# Comparative secretome analysis of normal prostate and prostate cancer cell models.

**Euan Johnstone**

*Thesis presented in fulfilment of the requirements for the degree of Master of Science in the Faculty of Biochemistry at Stellenbosch University.*

Supervisor: Dr K-H Storbeck
Co-supervisor: Dr M Vlok
March 2017

# Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2017

# Verklaring

Deur hierdie tesis elektronies in te lewer, verklaar ek dat die geheel van die werk hierin vervat, my eie, oorspronklike werk is, dat ek die alleenouteur daarvan is (behalwe in die mate uitdruklik anders aangedui), dat reproduksie en publikasie daarvan deur die Universiteit Stellenbosch nie derdepartyregte sal skend nie en dat ek dit nie vantevore, in die geheel of gedeeltelik, ter verkryging van enige kwalifikasie aangebied het nie.

Datum: Maart 2017

# Summary

Prostate cancer is the second most common cancer among men. The prostate specific antigen (PSA) was the first biomarker identified for the diagnosis of this prevalent disease. Although this biomarker is in routine use it has been extensively criticised and discredited due to the large number of false positives identified from its usage, which result from benign conditions such as benign prostatic hyperplasia and prostatitis. Several studies have therefore made use of genomic, transcriptomic and proteomic methods to identify novel candidate biomarkers for prostate cancer, which can replace PSA. The aim of this study was to contribute to this search by using state-of-the-art mass spectrometry based proteomics to characterise the proteome and secretome of benign (BPH-1), cancerous (LNCaP and PC-3) and normal (PNT2C2) prostate cell lines. The seeding densities of the four prostate cell lines were optimised for maximum protein secretion and reduced cell death in a chemically defined medium, with lower seeding densities yielding the best results. Acetone precipitation was subsequently found to yield the best protein recoveries from the conditioned media when compared to three other methods, which included ammonium sulphate, methanol chloroform and PTA-mediated acetone precipitation. The proteome and secretome profiles of the four cell lines were subsequently characterised by mass spectrometer based proteomics. A total of 3576 and 1106 proteins were positively assigned from the proteome and secretome samples, respectively. This data was subsequently analysed using Ingenuity Pathway Analysis, several upregulated molecular pathways, upstream regulators, molecular and cellular processes, disease states and potential biomarkers were identified. The data showed that pathways involved in the regulation of a Adherens junctions, oxidative phosphorylation and mitochondria were significantly upregulated in the prostate cancer cell lines and may therefore be useful avenues to pursue when searching for candidate biomarkers or therapeutic targets for prostate cancer. Furthermore, a total of 157 candidate biomarkers which could distinguish the prostate cancer cell lines from the benign prostatic hyperplasia cell line were identified. Despite a number of limitations this study further demonstrates the use of secretome based proteomics as a tool to complement biomarker discovery.

# Opsomming

Prostaatkanker is die tweede mees algemene kanker onder mans. Die prostaat-spesifieke antigeen (PSA) was die eerste biomerker geïdentifiseer vir die diagnose van hierdie algemene siekte. Alhoewel hierdie biomerker in roetinegebruik is, is dit gekritiseer en gediskrediteer as 'n biomarker vir prostaatkanker as gevolg van die groot aantal vals positiewes wat geïdentifiseer is met die gebruik daarvan, wat as gevolg van toestande soos nie-kwaadaardige prostaat hiperplasie en prostatitis is. Verskeie studies het dus gebruik gemaak van genomiese, transkriptomiese en proteomiese metodes om nuwe kandidaat biomerkers vir prostaatkanker, wat PSA kan vervang, te identifiseer. Die doel van hierdie studie was om by te dra tot hierdie soektog deur die gebruik van massaspektrometrie-gebaseerde proteomika. Die proteoom en secretoom van nie-kwaadaardige (BPH-1), kanker (LNCaP en PC-3) en normale (PNT2C2) prostaat sellyn was ontleed. Die lotingdigthede van die vier prostaat sellyne is geoptimiseer vir maksimum proteïen sekresie en minimale seldood. 'n Laer lotingdigthede het die beste resultate gelewer. Asetoonpresipitasie het die beste proteïenherwinning van die gekondisioneerde media gelewer. Die ander metodes wat getoets was, was die ammoniumsulfaat, metanol chloroform en PTA-bemiddelde asetoonpresipitasie metodes. Die proteoom en secretoom profiele van die vier sellyne is daarna met behulp van massaspektrometer bemiddelde proteomika ontleed. 'n Totaal van 3576 en 1106 proteïene is positief uit die proteoom en secretome monsters onderskeidelik, geïdentifiseer. Hierdie data is verder ontleed met behulp van "Ingenuity Pathway Analysis" sagteware. Verskeie veranderinge in molekuêre paaie, stroomop reguleerders, molekuêre en sellulêre prosesse, siektetoestande en potensile biomerkers is geïdentifiseer. Die data het getoon dat paaie wat betrokke is by die regulering van 'n Adherens aansluitings, oksidatiewe fosforilering en mitochondria is opgereguleer in die prostaatkanker sellyne en is dus nuttig paaie om te ondersoek vir kandidaatbiomerkers of terapeutiese middels vir prostaatkanker. Verder is 'n totaal van 157 kandidate biomerkers wat die prostaatkanker sel lyne van die nie-kwaadaardige prostaat hiperplasie sellyn kan onderskei, geïdentifiseer. Ten spyte van 'n aantal beperkings in hierdie studie toon die resultate dat sekretoom-gebaseerde proteomika gebruik kan word om biomerker ontdekking aan te vul.

# Dedication

*Dedicated to my friends and family who were always there and never let my love for science falter even when the times got tough.*

# Acknowledgments

I would like to thank following people and institutions that had a profound impact on the initiation, progression and completion of this project.

**Dr Karl-Heinz Storbeck** who without him none of this would be possible. Even though times were tough and there were some minor setbacks we managed to get there. Thank you for your constant support and guidance that allowed me to succeed in not only completing this project but also gaining a deeper understanding and passion for biochemistry.

**Dr Máre Vlok** for his dedication, technical prowess and editing skills that made not only the proteomic work possible, but exciting.

**Ralie Louw** for all the training, technical assistance and supplying me with everything I needed.

**Fellow students** including Kevin Muvirimi as a valued colleague, Jonathan Quanson for valued insights and entertainment, Monique Barnard and Elzette Pretorius for all their support, Stefan Hayward for his protein biochemistry insight, Timo Tait for his valued advice and Liezl Bloem for being a constant source of entertainment.

**My friends** especially Lauren Bezuidenhout, Kate Kuilman, Richard Atkinson and William Tipping-Woods who listened to me when I needed them most even though they never really knew what I was talking about.

**My family** who constantly supported me both financially and emotionally from start to finish.

**Stellenbosch University** for the space and environment necessary to succeed.

# Table of Contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| AR | Androgen receptor |
| ATF | Activating transcription factor 4 |
| BPH | Benign prostatic hyperplasia |
| *BRAF* | B-raf proto-oncogene, serine/threonine kinase |
| *BRCA2* | BRCA2, DNA repair associated |
| BSA | Bovine serum albumin |
| CAV1 | Calcium-sensing g-protein-coupled receptor, caveolin-1 |
| CDCHO | Chemically defined Chinese hamster ovary media |
| CID | Collision-induced dissociation |
| CRPC | Castration resistant prostate cancer |
| CTNNB1 | Catenin beta 1 |
| DHT | Dihydrotestosterone |
| DRE | Digital rectal examination |
| *ERG* | ETS Transcription Factor |
| ESI | Electrospray ionisation |
| ETS | ETS transcription family |
| ETV4 | ETS variant 4 |
| EZH2 | Histone-lysine N-methyltransferase EZH2 |
| FA | Formic acid |
| FBS | Fetal bovine serum |
| FDR | False discovery rate |
| FWHM | Full width at half maximum |
| GRK | G protein-coupled receptor kinases |
| *GSTP1* | Glutathione s-transferase $\pi$ |
| HAMSF12K | Ham's F12 Kaighn's modified media |
| HCD | Higher-energy collisional dissociation |
| HDI | Human development index |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HGPIN | High-grade prostatic intraepithelial neoplasia |
| IPA | Ingenuity pathway analysis |
| LC-MS | Liquid chromatography mass spectrometry |
| LDH | Lactate dehydrogenase |
| LFDR | Local false discovery rate |
| LGPIN | Low-grade prostatic intraepithelial neoplasia |
| LNCaP | Lymph node carcinoma of the prostate |
| $m/z$ ratio | Mass to charge ratio |
| MALDI | Matrix assisted laser desorption ionisation |
| mCRPC | Metastatic castration resistant prostate cancer |

| | |
|---|---|
| $mH_2O$ | MilliQ Water |
| MKI67 | Proliferation marker protein Ki-67 |
| MMTS | Methane methylthiosulphonate |
| MS | Mass spectrometry |
| MS/MS | Tandem mass spectrometry |
| MSI | Mass selective instability |
| NAD+ | Nicotinamide adenine dinucleotide (reduced) |
| NADH | Nicotinamide adenine dinucleotide (oxidised) |
| OGP | Octylgluco-mano-pyranoside |
| PBS | Phosphate buffered saline |
| PC-3 | Human prostatic carcinoma cell line |
| PCa | Prostate cancer |
| PenStrep | Penicillin-Streptomycin |
| PHI | Prostate health index |
| PI3K/AKT | phosphoinositide-3-kinase |
| PIK3CA | phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha |
| PIK3CB | Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta |
| PIN | Prostatic intraepithelial neoplasia |
| PNT2C2 | Normal prostate epithelium cells immortalised with SV40 |
| PRMT | Protein arginine methyltransferases |
| PSA | Prostate specific antigen |
| PSM | Peptide spectrum matching |
| PTA | Phosphotungstic acid |
| PTEN | Phosphatase and tensin homolog |
| *PTEN* | Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN |
| Q-TOF | Quadrupole time of flight |
| RF | Radio frequency |
| ROS | Reactive oxygen species |
| RPMI1640 | Roswell park memorial institute 1640 media |
| RSPO2 | R-spondin 2 |
| SRD5A | Steroid 5$\alpha$-reductase |
| TCEP | Triscarboxyethyl phosphine |
| TEAB | Triethylammonium bicarbonate |
| TFA | Trifluoroacetic acid |
| TMPRSS2 | Transmembrane protease serine 2 |
| TOF | Time of flight |
| Trypsin-EDTA | Trypsin-ethylenediaminetetraacetic acid |

| | |
|---|---|
| UniProt | Universal protein resource |

# Chapter 1

# An introduction to prostate cancer and biomarkers

## 1.1 Cancer

### 1.1.1 Cancer fact sheet

Cancer is one of the leading causes of non-communicable deaths worldwide.[1] Cancer is the collective term for a group of neoplasmic diseases characterised by the uncontrolled proliferation of cells with the ability to invade surrounding tissue or to metastasise.[2] This is in contrast to benign tumours which also result from the uncontrolled proliferation of cells, but lack these characteristics and are thus considered non-cancerous.[3]

Cancers can further be differentiated into five classes based on their origin: carcinomas (epithelial), sarcomas (connective tissue), lymphomas (blood or lymph), germ cells (pluripotent) and blastomas (embryonic).[4] Carcinomas are those that originate from epithelial tissues and account for more than 90% of all cancers.[3] Common carcinomas include prostate, colon, breast and lung cancers.[3] These cancers are usually age or environmentally related and are classic late onset cancers.[3] Sarcomas originate from mesenchymal tissues such as bone, connective tissue and cartilage and are prevalent in adults but occur at a higher rate in children.[5] Lymphomas originate in blood or lymph and include Hodgkin's and non-Hodgkin's lymphoma.[6] Germ cell cancers occur in the germ cells and are usually the result of birth defects or problems occurring during embryonic development.[7] Blastomas occur as a result of malignancies in precursor cells usually as a result embryonic developmental errors.[8]

Hormone dependent cancers are dependent on specific sex hormones and include breast, cervical, uterine, endometrial, testicle and prostate cancers.[9] The sex hormones are divided into androgens, oestrogens and progestogens which are all

steroid hormones produced in the gonads as well as in the adrenal glands, albeit in lower quantities.[10] Research has also shown that the presence of exogenous hormones from foods, contraceptives and hormone replacement therapies has exacerbated the incidence of hormone dependent cancers.[9]

While cancers may have many origins they all share the same traits, namely: unregulated cell proliferation caused by an increase in pro-proliferation signals or the reduction in anti-proliferation signals; the decreased production of suppressor genes; the reduction in controlled apoptosis; limitless proliferation potential; the ability to stimulate construction of blood vessels; the ability to invade tissues; and the ability to spread or metastasise.[11]

While cancer is generally attributed to the uncontrolled proliferation of cells and the development of tumours in most cases there are little to no early overt symptomatic traits that allude to its specific occurrence.[12] The non-specific nature of cancer means it often goes undetected or is misdiagnosed in patients until more severe symptoms require closer evaluation.[13] The development of tumours in or around specific organs do eventually indicate a disease state.[14] For instance in the case of lung cancer patients may present with a non-specific cough[15] and in the case of prostate cancer urinary incontinence.[16] These symptoms necessitate further testing which often leads to diagnosis. Currently large amounts of research is focused on better identifying the symptomatic traits that may aid in the early diagnosis of cancer.[17,18]

### 1.1.2 Global prevalence

Currently there are over 100 types of cancer globally with incidences and mortalities specific to certain regions often correlated to the human development index (HDI) status of the area.[19] The development index has direct correlation to quality of life and subsequent lifestyle traits.[19] Factors of the HDI that contribute to cancer incidence and mortality include: age, drug use, diet, obesity, exposure to cancer causing agents (carcinogens), cancer causing infections and genetic risk factors.[19] Cancer is the leading cause of mortality in countries with high to very high development indices while, cancer is second to communicable diseases in countries with medium and low development indices as seen in figure 1.1.[20] That being said non-communicable diseases are steadily becoming more and more of a burden on countries with a low to medium HDI.[21] In 2012 there was an incidence of 14.1 million cases of cancer globally resulting in 8.2 million deaths and accounting for 14.6% of the total deaths recorded that year.[1] While incidence in high to very high HDI was higher (56%) compared to low to medium HDI countries (44%) the mortality rates were higher in low to medium HDI countries (53%).[19] This can be attributed to the availability of treatment in these countries.[22] Certain cancers are more prevalent globally than other cancers with lung, breast, colorectal, prostate

and stomach cancers contributing to over 49% of the overall cancers.[23] The most common cancer in women is breast cancer (25.2%), while lung cancer (16.7%) is the most common cancer in men with prostate cancer coming a close second at 14.8%.[22,23]

The global cost of cancer places large burdens on economies not only from a treatment perspective but loss of incomes due to loss of life and was estimated to cost the global market 1.16 trillion US dollars in 2010.[1]

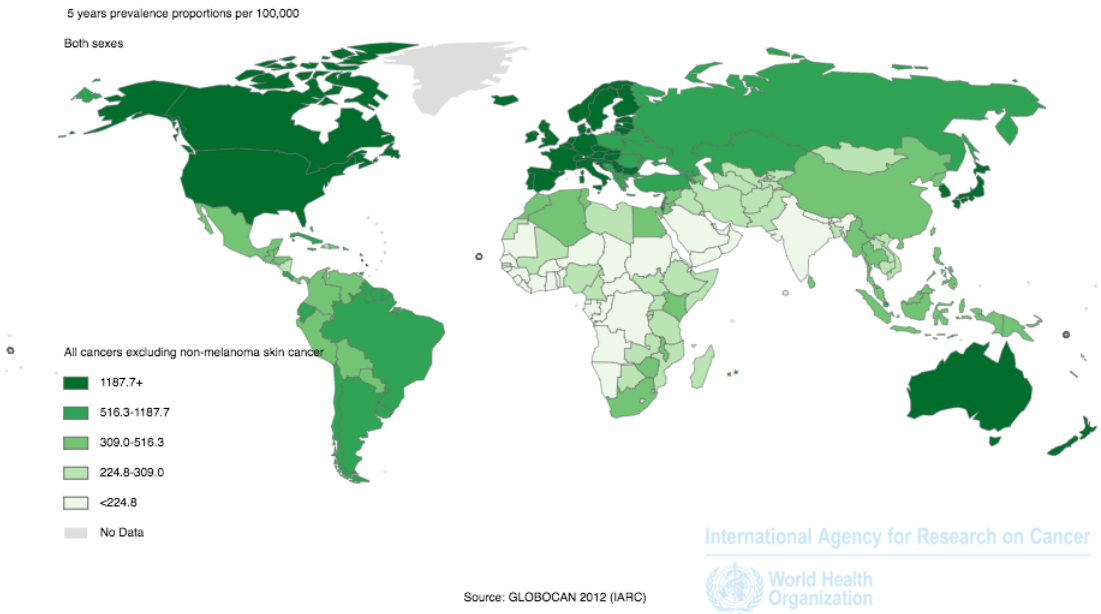


Figure 1.1: The global prevalence of all cancers (excluding non-melanoma skin cancer). Incidence is greatly increased in countries with high HDI compared to countries with low to medium HDI. Incidence was recorded as 5 year prevalence proportions per 100 000. Reproduced from GLOBOCAN 2012 (IARC).[19]

## 1.2 The prostate

The prostate is a small, internal, compound exocrine gland found in men located below the bladder on the anterior side of the rectum.[24] Its size and weight is dependent on overall body mass index of the individual but is usually referred to as walnut sized (3 cm long, 4 cm wide, 2 cm thick) weighing on average 20 grams.[25]

### 1.2.1 Prostate morphology

The prostate is a complex set of glands made up of distinct regions.[26] These regions were originally characterised anatomically as a set of four connected lobes including: the anterior lobe (isthmus), posterior lobe, median lobe or middle lobe and lateral lobes.[27] The prostate is now better characterised into three distinct glandular zones: the transition zone, the central zone and the peripheral zone as well as the anterior fibromuscular stroma as seen in figure 1.2.[28–30] The innermost transition zone directly surrounds the proximal urethra and accounts for 5-10% (5% at puberty) of the prostate.[26,30] It separates the bladder from the membranous urethra and is made up of pseudostratified columnar epithelium and secretory cells.[25,29] It is the only part of the gland that grows throughout a males lifetime and thus results in late onset urinary conditions such as benign prostatic hyperplasia (BPH).[31] The central zone is anterior to the transition zone and surrounds the ejaculatory ducts and makes up 25% of the prostate.[26,30] The central zone contributes to 2.5% of cancers that are considered to be more aggressive and invasive than cancers that affect other regions of the prostate.[32,33] The peripheral zone is posterior to the transition zone and makes up 70% of the prostate.[26,30] The peripheral zone forms the sub-capsular portion and surrounds the distal urethra of the prostate[32] and is the site of most prostate cancers.[34] The anterior fibromuscular stroma contains no glandular components and composed of muscle and fibrous tissue and extends between the urethra and transition zone.[35]

The prostate does not have a true capsule but is supported by randomly orientated fibromuscular stroma bundles.[35] The muscular component condenses peripherally to form the prostatic sheath and centrally merges into the muscle layers of the prostatic urethra.[36]

The prostate contains multiple gland types made up of various tissues.[25] The tubuloalveolar glands have pseudostratified columnar epithelium that vary in height.[25] The columnar cells, separated by basal cells contain apical secretory granules, while the basal cells are responsible for epithelium regeneration.[25] The prostate contains branching secretory tubules and acini that are surrounded by fibromuscular stroma.[37] In total the stroma accounts for about 70% of the prostate mass.[25]

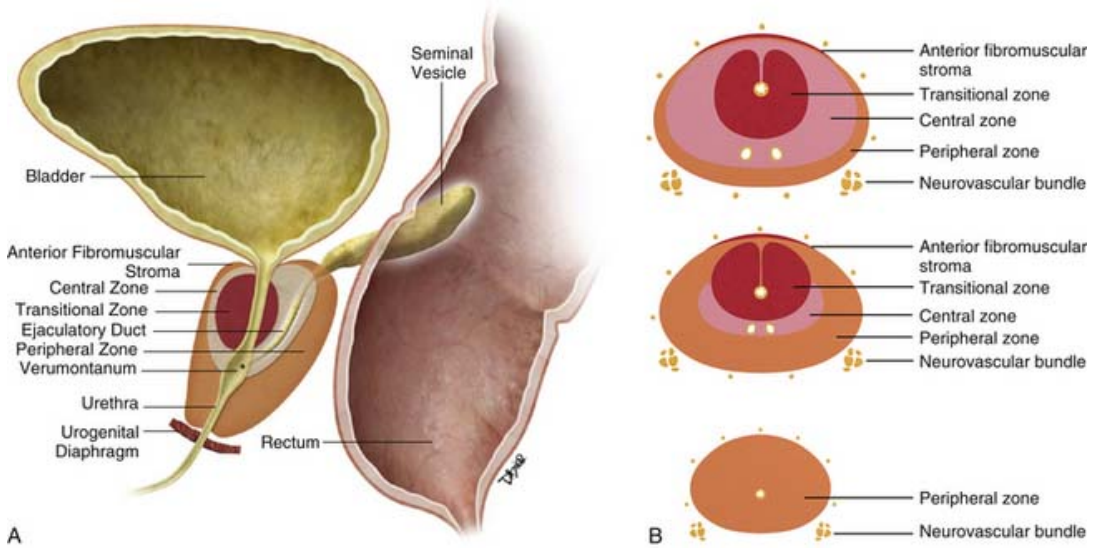


Figure 1.2: The relative location and zonal morphology of the prostate. The prostate contains the central zone, transitional zone and the peripheral zone surrounded by a fibromuscular stroma. Panel A depicts the sagittal view and panel B depicts the axial view with the top, middle and bottom sections depicting the base, mid gland and apex respectively. Reproduced from http://clinicalgate.com/prostate-cancer-4/.

### 1.2.2 Prostate function

The prostate is common among mammals and is referred to as an accessory sex gland.[25] It is not a critical part of the male reproductive but aids in fertility and urinary systems producing and combing constituents of seminal fluid and physical control of the urinary system.[38]

The central zone links the ductus deferens to the ejaculatory tubes and drains into the urethra.[32] This is the common site where the semen, fluid from the seminal vesicles and secretions from the bulbourethral gland are combined.[38] The prostate provides a variety of ions, lipids, simple sugars and a small portion of proteins (1%) to the ejaculate.[39] Some of these proteins include: proteolytic enzymes, prostatic acid phosphatase, beta-microseminoprotein and the prostate-specific antigen (PSA).[39,40]

PSA is produced by the secretory epithelium of the prostate and is responsible for the liquefaction of semen through the cleaving of semenogelin that increasing motility.[41,42] Upon damage to basal membrane and basal epithelial layer PSA leaks through the basal cells into the stroma and vasculature.[41,42] Serum PSA is thus elevated in patients presenting with prostate cancer and other prostatic conditions

such as BPH and prostatitis.[41,42] PSA is therefore widely used as a biomarker.[40]

The prostate also contains various muscular structures that are required to force ejaculate out of the urethra as well as prevent flow of urine into the seminal system.[38] The prostate and bladder sphincters work together to regulate the flow of urine and semen.[38] During ejaculation the internal sphincter closes to the bladder preventing flow black of semen into the bladder.[38] Alternatively during urination the seminal ducts close to prevent urine from entering the seminal system.[38] The prostate utilises smooth muscle fibres that help expel the fluid during ejaculation.[25] During ejaculation the prostatic acini are compressed by the surrounding fibromuscular coat that empties their components into the posterior urethra.[43]

### 1.2.3 Prostate development and maintenance

The development and maintenance of the prostate requires nutrients, oxygen and androgens (male sex hormones), which control the balance of proliferation and apoptosis.[44] The most important of these hormones is the androgen testosterone that is produced in the Leydig cells of the testes.[45] Testosterone production is dually regulated by luteinizing hormone and luteinizing hormone releasing hormone.[46] Androgens such as testosterone and its more potent form dihydrotestosterone (DHT) act by the ligand dependent transcription factor, known as the androgen receptor (AR) which is part of the nuclear receptor family and has multiple signalling mechanisms.[47]

The unliganded AR is located in the cytoplasm and is bound to heat shock protein 70 and heat shock protein 90.[48] Upon ligand binding these dissociate and the AR forms a homodimer which rapidly translocates into the nucleus where the liganded complex binds to specific DNA sequences known as androgen response elements (AREs).[48] The subsequent binding of cofactors allows the regulation of AR regulated genes such as PSA.[48]

Loss of AR activation leads to low prostatic development and prostate tissue atrophy, while the subsequent re-introduction of androgens leads to tissue regeneration, growth and secretory function.[49] Patients presenting with steroid 5$\alpha$-reductase (SRD5A) deficiencies resulted in atrophied prostate tissue.[49–51] SRD5A is responsible for the conversion of testosterone into the more potent androgen DHT required for normal prostate function.[52]

### 1.2.4 Prostatic conditions

The prostate is highly susceptible to inflammatory diseases such as BPH and prostatitis which usually occur later in life.[52] Prostatitis is the inflammation of the prostate caused by increased age or bacterial infection.[53] It can take four forms including the rarer acute or chronic bacterial form that is treated using antibiotics

or the more common non-bacterial or chronic pelvic pain syndrome treated using antibiotics and surgical reduction.[53]

BPH is the enlargement of the prostate that leads to the progressive blocking of the urethra that can impair adequate urinary function.[54] This is often treated using medication or more complex surgical means.[54] BPH is considered unavoidable and prevention is currently being investigated but it has been shown that that changes in diet and moderate intake of alcohol may have a beneficial effect.[55] BPH is highly prevalent occurring in 88% of men over the age of 80[56,57] and readily treatable[56,57] whereas prostate cancer (PCa) is less prevalent but far more deadly.[58]

## 1.3 Prostate cancer

PCa is the 2nd most common cause of cancer internationally accounting for 14.8% of male cancer diagnosis annually.[23] As of 2016 over one million men[23] are living with prostate cancer and a further 180 000 will be diagnosed in the USA alone.[59] PCa has a fairly large mortality rate of 6.6%[23] with over 26 000 deaths estimated for 2016 in the USA.[59] It is more prevalent in high HDI countries but is rapidly increasing in low to middle HDI countries[19] attributed to of better and more regular screening techniques.[60] PCa is the most common cancer among South African men, accounting for 17.15% of cancers annually and carrying a lifetime risk of 1 in 27.[61]

### 1.3.1 Prostate cancer risk factors

PCa usually doesn't occur as a result of one specific cause but instead it is attributed to the accumulation of genetic errors exacerbated by age, ethnicity, genetics, lifestyle and diet.[62] Age is by far the biggest risk factor when considering PCa incidence[62] with the risk steadily increasing from 0.6% in men aged 35-44 to 37.6% in men aged 65-74 with the median age of diagnosis being 66 years of age.[59] Ethnicity also plays a large role with prostate cancer being more prevalent in white males[61] but the associated risk being higher in black males.[62] This has been attributed to genetic as well as lifestyle factors exemplified by the increase in PCa prevalence in African Americans who have immigrated to HDI countries like the USA.[62] Genetics of PCa are still not fully elucidated but it has been shown that the incidence of prostate increases substantially in families with a history of PCa.[63] If a close relative such as a father or brother is diagnosed with PCa the risk of being diagnosed with PCa increases by a factor of 2.35 and 3.14 respectively.[62] Lifestyle factors such as smoking, lack of physical activity, obesity and alcohol intake have been shown to increase the risk of prostate cancer.[61] Diet and its effect on PCa has been comprehensively studied and has yielded confounding results possibly implicating, high fat, high red meat, low vegetable diets in increasing the risk of PCa.[61] PCa is

not readily preventable, but lifestyle changes such as diet changes, avoiding tobacco, low alcohol intake and avoiding other carcinogens may significantly reduce the likelihood of PCa[64]

### 1.3.2 Development of prostate cancer

The development and progression of PCa is complex and poorly understood.[65] Multiple pathways and mechanisms have been proposed building on historical, histological and genetic data.[66] Recent evidence has implicated precancerous conditions in the initiation and progression of PCa.[66] Large amounts of evidence have implicated prostatic intraepithelial neoplasia (PIN) that is preceded by proliferative inflammatory atrophy as an intermediate in the development of PCa.[66]

PIN is characterised by a number of lesions forming on the prostate classified according to their overall morphology including: cribriform, flat, micropapilary and tufting.[66] PIN can either occur as low-grade (LGPIN) or high-grade (HGPIN) depending on the extent of the lesions.[66] It has been hypothesised that HGPIN may be the immediate precursor to PCa due to a number of shared commonalities with PCa such as: HGPIN occurring in the transition zone; HGPIN diagnosis preceding PCa diagnosis by ten years, multiple similar genetic changes; similar morphologies; similar altered differentiation factors and increased telomerase activity.[66] Furthermore, PIN has been shown not to be precursor of other prostatic conditions such as BPH and prostatitis.[66] While HGPIN is similar to PCa it does not result in degradation of the basement membrane or increase the abundance of PSA.[66] Advanced PCa can metastasise rapidly and move from the prostate to surrounding tissue.[67] The final location is regulated by the distant microenvironment and chemokine gradients[67] and PCa preferentially migrates to bone, lung and liver[68] tissue utilising lymphatic, haematogenous and contiguous local spread.[69]

### 1.3.3 Genetic changes in prostate cancer

While it is easy to think that all tumours possess the same mutations and that these all occur at the same time, this is not the case and multiple different types of tumour mutations and tumour mutation profiles exist and interact to compound the effect of loss of control of cell proliferation.[70,71] Developmental pathways are critical in HGPIN and PCa initiation and progression.[72] The most fundamental epigenetic changes occur as a result of chromatin remodelling, DNA methylation, histone acetylation, alteration of transcriptional mechanisms and fusion translocations.[72] These alterations can lead to gain of function as well as loss of function mutations.[72] This way of thinking has given rise to the idea that cancer cells illicit an almost Darwinian approach to survival.[73] This is done by producing their own growth

signals, reducing apoptosis, suppressing inhibition signals, angiogenesis and inducing limitless replication potential and thus better competition for resources.[73,74]

Multiple studies have implicated genetic abnormalities in the progression of PCa to castration resistant PCa (CRPC) and progression to the more aggressive metastatic CRPC (mCRPC) and have elucidated possible avenues for more effective or actionable treatments.[75] The main genetic alterations which occur during the progression of PCa to mCRPC were identified during a recent study investigating the genetic landscape of PCa are shown in figure 1.3 below.[75] Genetic alterations were shown to result from single nucleotide variants, fusions, homozygous deletions and amplifications.[75] These alterations resulted in significant changes to multiple pathways including AR signalling, phosphoinositide-3-kinase (PI3K/AKT), WNT, DNA repair, cell cycle and gene modifier pathways.[75]

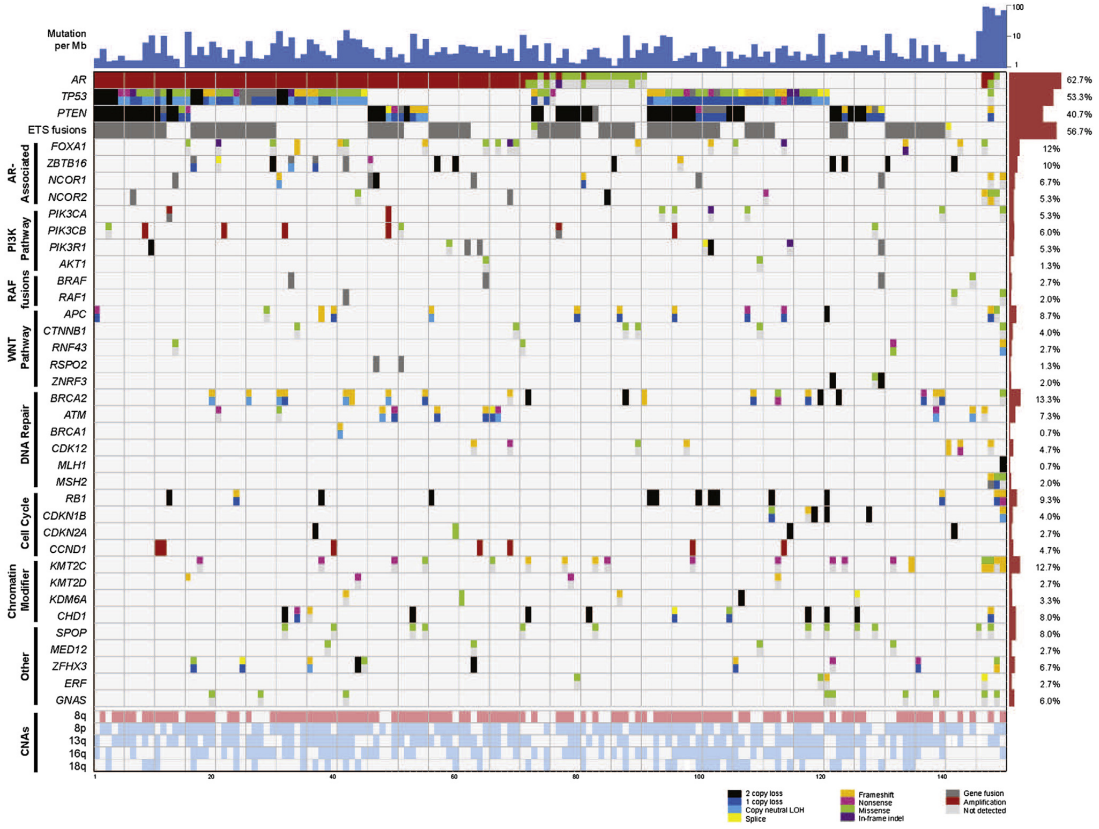


Figure 1.3: Integrative landscape analysis of somatic and germline aberrations in metastatic CRPC obtained through DNA and RNA sequencing of clinically obtained biopsies. Genes and their subsequent proteins can be further analysed to determine their effect on PCa. Reproduced from Robinson, D. et al. Integrative clinical genomics of advanced prostate cancer. Cell 161, 1215-1228 (2015).[75]

Chromosomal fusions were found to be a frequent feature of the various stages of PCa with an average of 15 gene fusions per tumour being detected.[75] These were mostly made up of ETS transcription family (ETS) to ERG, ETS transcription factor (ERG) fusions with less common fusions between ETS and Fli-1 proto-oncogene, ETS transcription factor (FLI1), ETS variant 4 (ETV4) and ETS variant 5 (ETV5).[75] Furthermore, fusions between ETS and B-Raf proto-oncogene, serine/threonine kinase (BRAF), phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA), phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta (PIK3CB) and R-spondin 2 (RSPO2) were also identified as potentially responsive to targeted therapy.[75] The BRAF, PIK3CA/B and RSPO2 act through regulation of the MAPK, PI3K/AKT and WNT pathways respectively.[76–78]

Chromosomal fusions between transmembrane protease, serine 2 (TMPRSS2) and ERG have also been shown to be highly prevalent.[72] These involve ETS transcription factors and TMPRSS2 and usually result in changes in the PI3K/AKT pathway.[72] AR signalling has been shown to induce TMPRSS2 and ERG proximity.[72] Proximity combined with recruitment of enzymatic cytidine deaminase and the long interspersed element-1 (LINE-1) repeat encoded open reading frame 2 (ORF2) endonuclease that results in double stranded DNA breaks and subsequent fusion through ligation by non-homologous end joining.[72] TMPRSS2 is androgen-inducible and is expressed in normal and PCa cells.[79] Therefore the fusion of TMPRSS2 with other genes allows them to become androgen inducible.[79] The final protein products of these fusions only include amino acids from the ERG gene indicating the TMPRSS2 fusion region acts as the androgen inducible promoter.[79] Thus the subsequent fusions allow upregulation of ERG, ETV1 and ETV4 in the presence of androgens.[79] ERG gene products regulate cell proliferation, differentiation and apoptosis[79] and TMPRSS2-ERG fusions therefore results in activation of oncogenic pathways.[72] TMPRSS2-ERG fusions also induce repressive epigenetic programs upregulating Histone-lysine N-methyltransferase EZH2 (EZH2).[72] ERG and the AR work together and inhibit AR mediated differentiation and promote EZH2 mediated dedifferentiation.[72] High levels of ETS in fusion positive cases also results in activation of RAS-responsive elements replacing the RAS-MAPK pathway.[72]

Multiple individual pathways have been shown to affect PCa development.[75] Many of these pathways and their constant interplay been shown to contribute to the onset and development of PCa as indicated by figure 1.4.[80] Alteration within the AR signalling pathway were shown to be common.[75] The main changes occur within the AR itself, truncated or missense mutations in the forkhead box A1 (FOXA1) pioneer transcription factor, disruption of the nuclear receptor corepressor 1 (NCOR1) and nuclear receptor corepressor 2 (NCOR2) negative regulators, mutations of

the putative transcription regulator speckle type BTB/POZ protein (SPOP) and deletions of zinc finger And BTB domain containing 16 (ZBTB16).[75] These result in an upregulation of the MAPK pathway as well as hotspot mutations in residues on the AR that confer agonism to AR agonists.[75] The hepatocyte nuclear factor 3-alpha (FOXA1) pioneer transcription factor is critical in the growth and differentiation of prostate cells and has been shown to negatively inhibit the activity of the AR and the PI3K/AKT pathway and thus has been implicated as a tumour suppressor.[81] Zinc finger and BTB domain-containing protein 16 (ZBTB16) negatively regulates transcription through binding to DNA and recruitment of co-repressors and has been shown to suppress AR activity. Loss of function leads to tumour proliferation as well as AR independence.[82]

The PI3K/AKT pathway was also shown to be significantly altered in many cases, primarily as a result of biallelic loss of phosphatase and tensin homolog (PTEN) as well as hotspot mutations, amplifications and activating fusions in PIK3CA resulting in amplification and overexpression and the p.E17K activation mutation which leads to PI3K-independent activation of serine/threonine kinase 1 (AKT1).[75,83] There were also fusions of the PIK3CA and PIK3CB resulting in overexpression increasing the activity of the PI3K/AKT pathway which is responsible for cellular growth and proliferation.[75,84] Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN (PTEN) phosphorylates downstream targets such as phosphatidylinositol (3,4,5)-trisphosphate (PIP3) that results in the inactivation of the PI3K/AKT signalling pathway and is a known tumour suppressor.[72] While PTEN mutations are rare, expression has been found to be greatly reduced resulting in the increased activity of the PI3K/AKT pathway that has been shown to increase cellular resistance to apoptosis.[66]

The WNT pathway is responsible for the direction of embryonic development having specific effects on stem cell specification, proliferation, polarity, migration and maintenance.[80] Its interaction with the PI3K/AKT pathway, AR and proteins such as PTEN and catenin beta 1 (CTNNB1) has been shown to contribute to PCa tumourigenesis.[80] Alterations in the WNT pathway were shown to result from hotspot activations of CTNNB1, and recurrent alterations of APC WNT signalling pathway regulator (APC). Furthermore, alterations in ring finger protein 43 (RNF43), and zinc and ring finger 3 (ZNRF3) were mutually exclusive to APC mutations.[75] Additionally RSPO2 fusions resulted in its overexpression.[75]

It is well know that PCa occurs as a result of changes in the cell cycle.[75] The role of retinoblastoma transcriptional corepressor 1 (RB1) was implicated in these changes as well amplifications of cyclin D1 (CCND1) and less commonly cyclin dependent kinase inhibitor 2A/B (CDKN2A/B), cyclin dependent kinase inhibitor 1B (CDKN1B) and cyclin dependent kinase 4 (CDK4).[75] Retinoblastoma-

associated protein (RB1) negatively regulates proliferation through mediation of promoter regions of cyclin A2 (CCNA2) and minichromosome maintenance complex component 7 (MCM7) as well as inactivation of E2F-mediated transcription.[85]

Changes in the DNA repair pathway have significant effects on PCa progression with alterations of BRCA2, DNA repair associated (BRCA2).[75] Breast cancer type 2 susceptibility protein (BRCA2) is a known tumour suppressor responsible for mitigating DNA damage by efficient and precise repair of double-stranded DNA breaks[86] Biallelic loss of BRCA2 occurred as a result of somatic point mutations, loss of heterozygosity and homozygous deletion.[75] BRCA2 loss may also occur as result of pathogenic germline mutations combined with somatic alterations resulting in biallelic loss.[75] There was also a large prevalence of biallelic loss of ATM serine/threonine kinase (ATM).[75] Alterations were also shown in BRCA1, DNA repair associated (BRCA1), cyclin dependent kinase 12 (CDK12), Fanconi anaemia complementation group A (FANCA), RAD51 paralog B (RAD51B), RAD51 paralog C (RAD51C) as well as hypermutation events occurring as a result of structural changes in mutL homolog 1 (MLH1) or mutS homolog 2 (MSH2).[75]

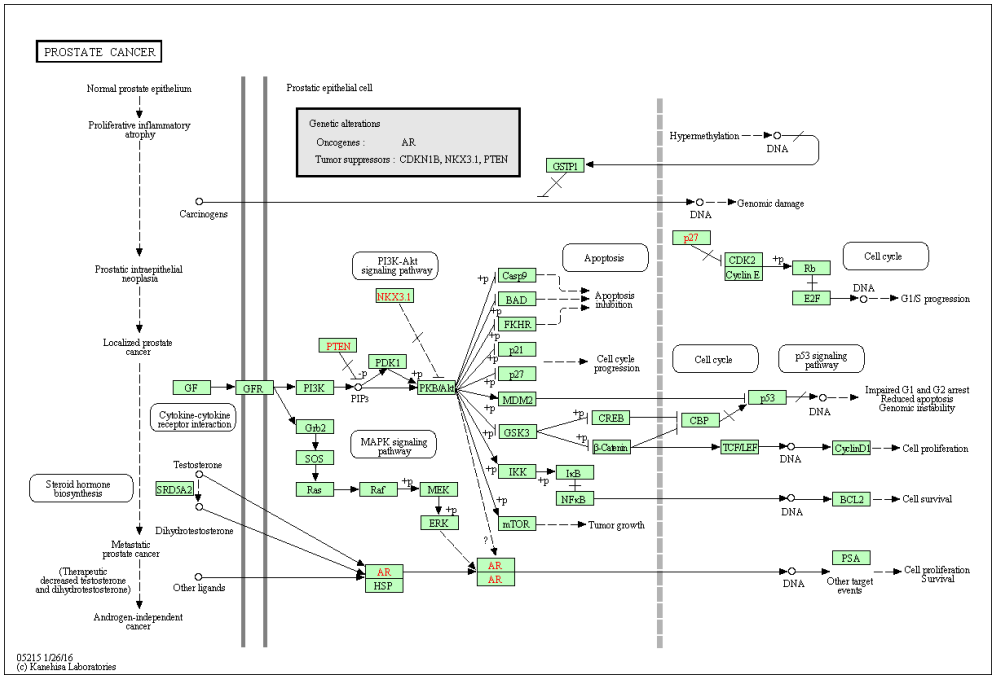


Figure 1.4: A Kyoto Encyclopaedia of Genes and Genomes (KEGG) representation of the possible pathways implicated in the development and progression of PCa. Protein upregulation and downregulation events can be analysed to determine their individual effects on pathways. Reproduced from http://www.genome.jp/kegg-bin/show_pathway?hsa05215

### 1.3.4 Mortality

Cancer when detected early is usually treatable.[3] The time taken to discover the presence of cancer is directly linked to the final prognosis.[87] Once the cancer has metastasised it becomes far harder to treat and manage.[87] The usual progression of cancer is, pre-malignant, primary tumourigenesis, rapid genetic alteration, metastasis and invasion.[88] Once the cancer has fully spread treatment is still possible but the likelihood of full recovery is unlikely and treatment usually focuses on increasing quality of life or increasing life expectancy through the use of palliative care.[89] The progression from treatment failure to death is dependent on the spread and grade of the cancer.[90] This time period can run from weeks to years.[91] Although cancer is the overall cause of mortality the actual causes of death can be attributed to: infection, haemorrhages, thromboembolisms, organ invasion, cachexia, respiratory failure and cardiovascular failure.[92] It has been widely reported that patients with higher qualities of life tend to live longer.[93]

### 1.3.5 Prostate cancer treatment

After diagnosis and classification a patient prognosis is determined outlining the chances of survival and the overall seriousness of the cancer.[90] The prognosis is determined by a wide range of factors including the type and traits of the cancer.[90]

PCa treatment is highly controversial. Small benign tumours are primarily treated with either surgical reduction (radical prostatectomy) or radiation therapy.[94] Treatment is often unnecessary as tumours may have remained benign and treatment has been shown to increase prostate cancer mortality.[95] The next available treatment option is Brachytherapy through the insertion of radioactive implants directly into the prostate tissue.[96] The final treatment option for more aggressive cancers is androgen deprivation therapy (ADT) achieved through chemical or surgical castration.[97] Primary ADT utilises either a gonadotropin-releasing hormone analog or gonadotropin-releasing hormone with or without an oral anti-androgen resulting in combined androgen blockade therapy.[98] This results in the reduction of circulating androgens that has been shown to be highly effective resulting in large amounts of apoptosis.[99] However, in most cases the PCa eventually resurges as the more aggressive CRPC.[100] CRPC is treated using secondary ADT that utilises anti-androgens and inhibitors of androgen production.[101]

### 1.3.6 Prostate cancer diagnosis

PCa screening and diagnosis is steadily becoming more widespread.[102] Common practice screening methods include the digital rectal examination (DRE),[103] PSA level testing[103] and the prostate health index (PHI).[104] The DRE and PSA level

testing have been shown not to reduce the mortality of PCa.[105] Positive DRE and PSA results give no indication of the severity of PCa and require further testing.[106] The PHI is a mathematical model utilising PSA, free PSA and proPSA that increases the sensitivity of PCa diagnosis.[104] Further testing is highly invasive and includes a transrectal ultrasound (TRUS) and transrectal ultrasound-guided systemic biopsy[107] that can result in complications.[108] PSA has been shown to be highly non-specific and thus necessitates the need for the development of a new more effective diagnostic method.[106]

## 1.4 Biomarkers

### 1.4.1 Biomarker fact sheet

The term biomarker has multiple definitions.[109] The broadest definition is an objective indication of a medical state observed from outside the patient which can be measured accurately and reproducibly.[110] A more applicable definition is that of the National Cancer Institute; "A biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease. A biomarker may be used to see how well the body responds to a treatment for a disease or condition. Also called molecular marker and signature molecule."[109] The earliest example of a cancer biomarker was the discovery of light chain of immunoglobulin in urine in 1848 that was indicative of myeloma.[111] Since then multiple biomarkers have been identified for multiple conditions including PSA for prostate cancer which was identified in 1980.[111]

Biomarkers can come in a number of forms from the DNA to the metabolite level.[111] They can be characterised into three categories including, diagnostic, prognostic and stratification and validated in terms of specificity and sensitivity.[111] Specificity and sensitivity are often incorrectly used interchangeably.[112] Specificity refers to the portion of the population that are identified as true negatives while sensitivity refers to those that are identified as true positives.[111] Concurrently high levels of specificity and sensitivity are very hard to attain and receiver operating characteristic curves as demonstrated in figure 1.5 below are used to identify potential biomarkers of high enough specificity and sensitivity.[111] A diagnostic marker must be highly specific and sensitive and must indicate a disease state.[111] Prognostic markers are used after a disease state has been identified and is used predict the progression of a disease, its possible reoccurrence as well as the possible outcomes of treatment.[111] Stratification or predictive biomarkers are used to predict the effects of a treatment prior to treatment application that better help identify patients who will respond to which treatment.[111]

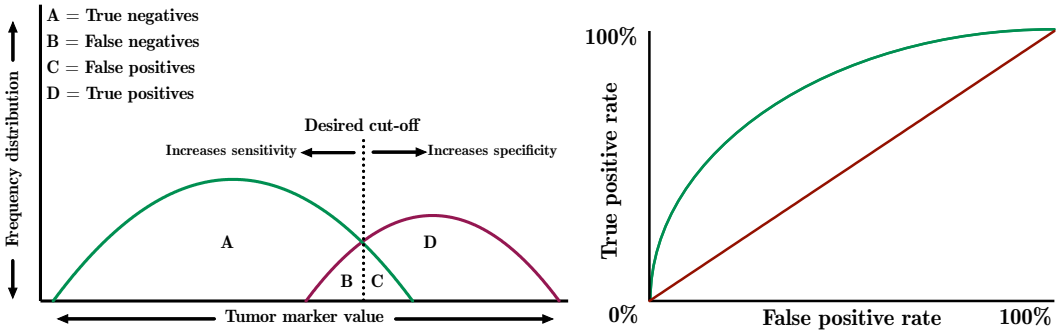


Figure 1.5: A receiver-operating curve representing the relationship between sensitivity and specificity used to evaluate the efficacy of a tumour marker at various cut-off points. Receiver operating curves can be used to identify the true positives, true negatives, false positives and false negatives. Reproduced from http://www.nature.com/nrclinonc/journal/v5/n10/full/ncponc1187.html

The ideal characteristics of a good biomarker include the ability to increase the survival rate or quality of life of a patient, ease of use, reliability and cost.[111] Currently most biomarkers require further tests to successfully identify a disease state.[111] Specifically in cancer biomarkers a number of characteristics are ideal and include: biomarkers produced by tumours and present in circulation; quantifiably present at low quantities in healthy or benign tumour individuals and high in diseased individuals; easily and cost effectively quantifiable; quantitatively relative measurement to tumour burden; and high sensitivity and specificity.[111]

### 1.4.2 Prostate specific antigen

PSA was first identified in 1971 and its use as a clinical biomarker developed for PCa in 1980.[113] Initially it was deemed a highly promising biomarker, but has been steadily discredited due to its low specificity and sensitivity.[106] While elevated levels of PSA were initially thought to be indicative of PCa it has been shown that elevated levels of PSA are also present in patients presenting with other prostatic conditions such as prostatitis and BPH.[114] PSA may also become elevated due to external prostate interactions such as the DRE or prostate biopsies.[114] PSA is therefore not indicative of PCa and has led to severe over-diagnosis often resulting in unnecessary and often harmful further diagnostic and treatment procedures.[106] Further unnecessary diagnosis may include core needle biopsies which themselves may not be effective and potentially harmful.[115]

The lack of sensitivity and specificity of PSA has necessitated the development for novel biomarkers and has yielded significant results.[106] Several new biomarkers

have been identified as potential diagnostic, prognostic and stratification biomarkers including, disabled homolog 2-interacting protein (DAB2IP), Golgi membrane protein 1 (GOLM1) proliferation marker protein Ki-67 (MKI67), *TMPRSS2-ERG* fusion, prostate cancer associated 3 (*PCA3*), PDZ and LIM domain protein 4 (*PDLIM4*), caveolin-1 (CAV1), sarcosine, chromogranin A (CHGA), glutathione s-transferase $\pi$ (*GSTP1*) and alpha-methylacyl-coa racemase (AMACR).[106]

MKI67 is a promising prognostic biomarker for PCa.[106] It has been shown to predict the outcome of a radical prostatectomy in patients who are considered low-stage or low grade PCa.[106] It may allow more effective treatment and targeted treatment of PCa.[106]

CAV1 is also potentially promising as it can differentiate between various prostatic conditions based on the relative levels.[106] Conditions such as BPH express a large amount of CAV1 while early stage PCa expressed lower levels and metastatic PCa showed no expression.[106] This means that CAV1 may be useful in targeted treatment and prognosis.[106]

*GSTP1* is a promising diagnostic biomarker for PCa.[106] The *GSTP1* gene has been shown to be unmethylated in normal prostate tissue and BPH tissue and hypermethylated in PCa tissue.[106] It is readily secreted into serum, urine and seminal plasma meaning it can be easily and quantitatively analysed and is indicative of early stage PCa and PIN.[106] This is potentially the most promising and useful biomarker for diagnosis and prognosis pre and post treatment.[106]

### 1.4.3 Biomarker discovery

Discovery of novel biomarkers falls into five distinct phases, preclinical exploratory studies, assay development and validation, retrospective longitudinal clinical repository studies, prospective screening studies and randomised control trials.[111] Preclinical exploratory studies identify biomarkers by investigating differences between normal and tumorigenic cells.[111] Assay development and validation focuses on using a positive identification biomarker to develop a test for positive cases only.[111] Retrospective longitudinal clinical repository studies then check if negative or early diagnosis can be achieved utilising the same biomarkers.[111] Prospective screening studies are the refinement of the assay to identify positive individuals as well as the extent of the disease state.[111] Finally randomised control trials are used to identify the population wide effect of using the novel biomarker.[111]

Biomarkers can be discovered at every level of 'omics, namely genomics, transcriptomics, proteomics, sectretomics, metabolomics and lipidomics.[116] These approaches are constantly becoming more powerful with modernisation and development of new technology.[111]

Genomics and proteomics are the two most utilised methodologies. Genomics utilises microarrays, northern blotting, standard gene expression techniques and

serial analysis of gene expression while proteomics is further downstream and utilises, mass spectrometry, 2D polyacrylamide gel electrophoresis, antibody arrays and tissue microarrays.[117]

A very promising approach for proteomic biomarker discovery and improvement is the use of tandem mass spectrometry.[111] This field has slowly transitioned from large solitary proteins to low molecular weight groups of proteins that can increase sensitivity and specificity.[111] Disease states and populations are highly heterogeneous leading to a later focus on more broad identification of candidate biomarkers.[111] The proteomic level has a distinct advantage of upstream methods such as genomics and transcriptomics as transcription and translation events do not directly translate into a defined product and may undergo multiple changes that may not directly indicate a disease state leading to the theory that proteins hold more physiological characteristics than genes.[118]

A subsection of proteomics is secretomics that is the study of proteins secreted into the surrounding environment such as serum or media.[119] Secretome analysis is becoming steadily more prominent for its success in identifying candidate biomarkers that are inherently non-invasive to obtain.[119] Secretomics samples are relatively easy to obtain and changes in methodologies has refined the technique significantly.[119] It is highly useful in cancer biomarker discovery as secreted tumorigenic proteins are readily involved in tumour progression.[119]

There are two strategies in proteomics, top-down and bottom-up.[120] Top-down is the utilisation of intact proteins prior to ionisation and detection whereas bottom-up proteomics utilises pre-digestion of proteins into peptides prior to ionisation.[120] In a bottom up approach samples typically undergo isolation techniques prior to digestion to increase the resolving power and add a level of information such as subcellular location.[121] Proteins are then digested either enzymatically or chemically.[121] Trypsin is a commonly used enzymatic digestion agent and cleaves proteins at lysine and arginine residues.[122] Peptides then typically undergo fractionation by liquid chromatography prior to MS analysis.[122] Liquid chromatography allows for the separation of peptides and can be coupled directly to the mass spectrometer, thus reducing the risk of sample loss.[122]

### 1.4.4 Mass spectrometry

Mass spectrometry (MS) is the study of simple or complex samples by ionising the components and sorting the resulting ions based on mass to charge ($m/z$) ratio.[123] The mass spectrometer was initially used to separate uranium ions during the Manhattan project.[124] A simple flow through of mass spectrometry involves: the ionisation of a sample, separation of subsequent ions by their $m/z$ ratio in a magnetic or electric field and detection.[120] In the case of proteins, ionisation is achieved by electrospray (ESI) or matrix assisted laser desorption ionisation

(MALDI).[123] MALDI is more suited to intact mass analysis than ESI (top down approach). This approach is, however, hampered by the complex fragmentation patterns which are difficult to interpret using current algorithms.[121] The majority of proteomic studies, therefore employ ESI of proteins which are first enzymatically digested into smaller peptide fragments (bottom up approach). ESI is then used to induce positive or negative charges on ions through protonation/deprotonation or cationisation.[123] The resulting peptide ions are then accelerated using a highly positively or negatively charged electric field and are then passed through various types of mass analysers which filter the ions based on $m/z$ ratio allowing only selected ions to reach the detector at any given time[123] A large variety of mass spectrometers are available each with different configurations. This study made use of an Oribitrap Fusion Tribrid mass spectrometer. These mass spectrometers are highly efficient and provide a high throughput platform for conducting bottom up proteomic studies.[125] The standard Tribrid architecture combines an ionisation source, quadrupole, dual pressure linear ion trap and an ultra-high-field Orbitrap mass analysers.[126] The individual components are discussed below.

### Electrospray ionisation

ESI based mass spectrometers can be coupled to liquid chromatography based separation systems. This allows for the separation of peptides prior to ionisation thus reduces the complexity of peptides entering the mass spectrometer at any given time.

ESI uses electrical energy to aid in the transfer of ions from liquid into the gas phase allowing the identification of ions in solution.[123] There are three parts to ESI, the production and dispersal of a fine spray of droplets, solvent evaporation and ion ejection from the highly charged droplets as depicted by figure 1.6.[123] These ions are ejected into a highly electrically charged tube with a polarity matched to the ions.[123] Furthermore, a nebulizing gas such as nitrogen is introduced to the chamber shearing the ionic droplets which increases and enhances the flow rate.[123] As the ions are accelerated out of the tube and prior to detection by a mass analyser a pressure and potential gradient is applied to aid in separation of the ions.[123] Furthermore, an increased temperature or the introduction of a nebulizing gas may be applied to reduce the amount of solvent present. This decreases the droplet radius and increases the surface charge density.[123] Eventually the electric field strength within the ions reaches its maximal limit that allows the ions to overcome the surface tension of the droplet and be ejected into the gas phase.[123,127] These ions then enter various types of mass analysers which filter the resulting ions based on $m/z$ ratio.[123]

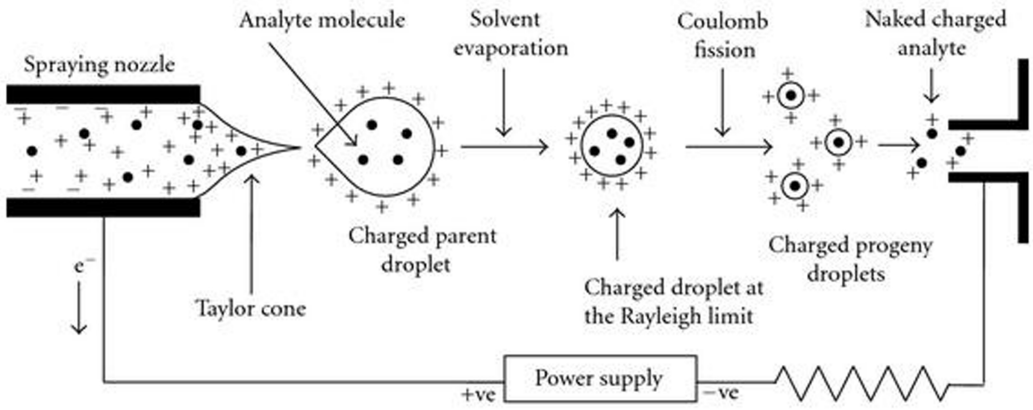


Figure 1.6: Schematic of the principles applied in ESI. Reproduced from http://lab-training.com/wp-content/uploads/2016/01/Schematic-of-Electrospray-ionization.jpg

### Mass analysers

**Quadrupole mass analyser** Quadrupole analysers allow easy coupling to ESI and are relatively small, robust, economical to operate.[123] They utilise the $m/z$ ratio as their primary means of separating and identifying ions.[123] Four parallel rods are placed equidistant from one another.[123] Equal DC voltages superimposed with a radio frequency (RF) voltage are applied to each set of diagonal rods that generates an electrical field that allows the ions to move forwards through the field in an oscillatory motion.[123] By setting the ratios of DC to RF voltage applied to the rods the ions travel at defined amplitudes that are related to defined $m/z$ ratios.[123] This allows the selection of ions at defined $m/z$ ratios to reach the detector or subsequent mass analysers and causes all other ions to collide with the rods neutralising them and preventing them from being detected.[123]

**Tandem quadrupole mass analyser** Quadrupole mass analysers can be placed in series resulting in a tandem quadrupole system, often denoted as MS2.[123] Typical tandem quadrupole systems contain three quadrupole analysers denoted as Q1, Q2 and Q3.[123] Q1 acts a pre-selection system selecting the ion of interest.[123] The ion of interest is then applied to Q2 which usually only utilises RF current and contains an inert gas such as argon.[123] The ions collide with the gas converting the kinetic energy to internal energy, become activated and are further fragmented into product ions using collision induced dissociation (CID).[123,128] The product ions

then enter Q3 where they are filtered by $m/z$ ratio.

**Ion trap mass analyser** Ion trap mass analysers became available after quadrupole mass analysers and are more economical, smaller and allow tandem CID monitoring.[123] They consists of three hyperbolic electrodes: the entrance end cap electrode, ring electrode and exit end cap electrode.[123] The combination of these allows the trapping and storing of ions between the entrance and exit cap allowing highly specific analysis.[123] Ions are usually pre-separated in a quadrupole mass analyser before entering the ion trap through a small hole in the entrance end cap electrode.[123] A voltage is then applied to the electrodes to trap ions. Ions can then be ejected at defined $m/z$ ratios by altering the voltage of the electrodes.[123] By applying a RF potential, an AC potential of constant frequency and variable amplitude to the ring electrode a three dimensional quadrupolar field is generated in the trapping area.[123] This field traps the ions in an oscillating trajectory dependent on the $m/z$ rations of ions in the trapping cell.[123] The ions are ejected from the trapping cell by inducing instabilities through alteration of the electrode system potentials.[123] These ions are ejected in increasing $m/z$ ratio in an axial direction from the oscillating trajectory and then detected.[123] Tandem MS can be used whereby the precursor ion is selected in the trap and an inert gas is applied inducing CID.[123] The product ions are then ejected and detected.[123] Product ions can be retained in the trap and subjected to multiple rounds of CID that aids in the differentiation of molecules that have similar structures.[123] Ion traps analysers are limited in their modes of acquisition and cannot utilise precursor or neutral loss scans therefore reducing their quantification potential by ten times when compared to tandem quadrupole systems.[123]

**Quadrupole Time of Flight mass analyser** Further advancements in mass spectrometry led to the development of the quadrupole time of flight mass spectrometer (Q-TOF).[129] The Q-TOF is similar to the triple quadrupole as it contains two quadrupoles with the addition of time of flight (TOF) mass spectrometer as depicted in figure 1.7.[129] An additional pre-quadrupole may be added prior to the two quadrupoles to aid in collisional dampening.[129] Ions are generated through a suitable ion source such as ESI or MALDI and travel through the initial RF only pre-quadrupole. Ions collide with a high-pressure nebulizing gas which focuses the ions by reducing energy spread and beam diameter and allows better transmission through the subsequent quadrupoles and TOF analyser.[129] Ions then travel through the first quadrupole, which, during MS/MS is run in mass filter mode to select a precursor ion of interest.[129] The precursor ions of interest are transmitted to the second quadrupole where they undergo fragmentation via CID.[129] The interaction of the ions with the collision gas further focuses the ions, which is critical for

TOF analysis.[129] Fragmentation may be avoided to obtain single mass spectra by significantly reducing the collision energy but this is not recommended, as it does not result in additional focussing.[129] Ions are once again accelerated and ion optics are used to focus the ions into the ion modulator portion of the TOF analyser.[129] The modulator region is initially field free and allows the ions to proceed in their original trajectory.[129] A pulsed electric field is applied across the modulator gap forcing the ions into an orthogonal trajectory and into an accelerating column where they obtain the necessary energy for detection.[129] The ions move into the field free region for TOF mass separation.[129] The ratio of energies is utilised to allow the ions to proceed to the mass detector without external electric interference, which may reduce the mass resolution.[129] A single stage mirror is used to correct the initial spatial spread of the ions.[129] Any ions that have deviated from the original vertical positions are focussed on a horizontal plane at the detector entrance.[129] Ions collide with two microchannel plates, are detected and recorded by a time to digital converter which counts the relative abundances of ions[129] Counting can be additionally improved through the use of a multiple anode detector to increase dynamic range.[129] The Q-TOF provides high mass accuracy, mass resolution and sensitivity.[129]

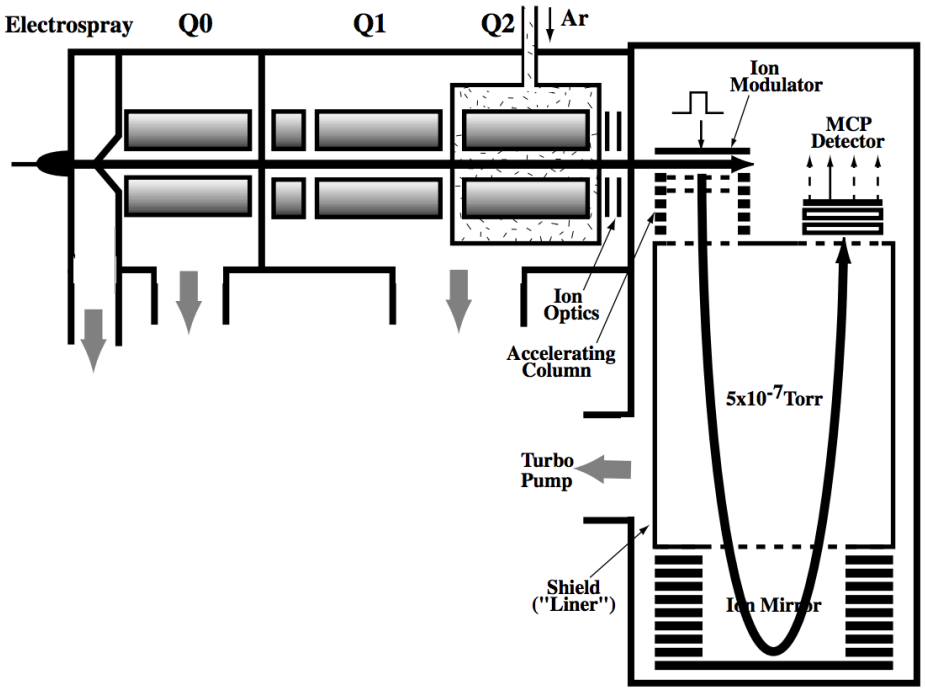


Figure 1.7: A quadrupole Time of Flight mass spectrometer. Reproduced from Igor V. Chernushevich,Alexander V. Loboda and Bruce A. Thomson.[129]

**Orbitrap mass analyser** Mass spectrometry methodology is constantly evolving requiring equipment capable of cost effective high performance data acquisition.[130] Alexander Makarov built on the principles of the Kingdon ion trap, quadrupole mass analyser and Fourier transform ion cyclotron resonance instrument to develop the Orbitrap mass spectrometer.[130] The Orbitrap is made up two electrodes: an outer-barrel like electrode responsible for excitation and detection of ions and a coaxial inner spindle-like electrode responsible for maintaining the ions orbital motion.[130,131] Active measurements are taken in the space between the electrodes.[125] Orbitrap mass analysers measure multiple consecutive small packets of ions in a short time frame and require coupled systems to allow continuous measurement of ions.[130] This was initially its major pitfall that was remedied through the use of a coupled system.[130] A coupled system usually involves and ionisation source such as ESI, three tandem linear quadrupole mass analysers, a lens system and lastly the Orbitrap itself.[130] Initially ions generated from ESI and travel into the first and second high-pressure quadrupoles where they are selected, guided and transported to a third low-pressure storage quadrupole.[125] As they enter the storage quadrupole the ions come into contact with a nebulizing gas that slows downs the ions and allows them to pool, accumulate and maintain their small spatial extent before they ejected into the Orbitrap in distinct packets.[124] Distinct packets improve the signal to noise ratio and resolution.[124] The architecture allows CID or higher-energy CID at any stage of MS.[125] Once enough ions have accumulated the exit lens of the storage quadrupole is opened generating a strong electric field that forces and accelerates the ions into the Orbitrap.[124] As they pass through the exit ring a small DC potential is applied that helps generate the axial motion utilised in the Orbitrap.[124] The ions enter the Orbitrap off centre to further increase the axial motion and ensure the ions are travelling coherently.[130,131] To further ensure the ions are travelling coherently a small voltage is applied to one part of the outer electrode that forces the ions into defined axial position and when the current is removed the ions move in a coherent motion.[130,131] At this stage the all the ions have axial, radial and rotational motion.[130,131] The ions possess identical axial oscillations and amplitudes but ions with different $m/z$ ratios will possess a different frequency.[124] The ions are trapped in the Orbitrap and cycle around and oscillate back and forth along the inner electrode.[124]

Mass analysis can be achieved through two modes; Fourier transform and mass selection instability (MSI) mode.[130,131] Fourier transform mode utilises the outer electrode to measure the coherent oscillations of the ions in the axial direction through image current detection.[130,131] The current image is then transformed using Fourier calculations that generate mass spectra.[130,131] Measurements taken using Fourier transformations have extremely high resolution.[130,131] MSI mode is similar to traditional detection systems.[130,131] Ions are ejected from the Orbitrap

that collide with a dynode that generates secondary electrons, which are multiplied, collected and detected.[130,131] MSI can be achieved using either paramagnetic resonance or resonance excitement.[125] Paramagnetic resonance utilises a voltage applied to the electrodes, which is varied sinusoidally with time.[125] In resonance excitation sinusoidally oscillations are applied to half of the outer electrode.[125] The advantage of MSI is that it allows storage of ejected ions for MS/Ms analysis and can act as a filter by ejecting contaminating ions, which improves the dynamic range.[125]

The Orbitrap reduces problems associated with normal quadrupole mass analysers and ion traps such as low mass accuracy, limited linear range, high complexity, and reduced charge capacity.[125] This architecture results in extremely high resolution, mass accuracy and dynamic range.[125]

### Mass spectrometry data acquisition

After mass analysis the ion data is converted to a graphical representation called a mass spectrum indicating the $m/z$ ratios and their relative abundances.[123] Quantitative information is slightly harder to obtain.[123] Utilising ESI-MS the ion signal detected is proportional to ion concentration and independent to flow rate and injection volume.[123] Therefore it is important to incorporate reference sample to account for loss accrued during sample preparation and variable detection during mass analysis.[123]

Intact protein mass spectrometry (top down approach) has distinct advantages and disadvantages. For instance information can be obtained on post-translational modifications, polymorphisms and sequence variants with relative ease but intact proteins are prone to developing multiple charge states, which undergo multiple different fragmentation patterns, which prevent the accurate determination of molecular masses.[121,132] A more common approach is to digest the proteins to peptide fragments prior to MS analysis (bottom up approach).[133]

The bottom up approach is an indirect method in terms of quantification as it is based exclusively on protein abundance, which is measured by spectral counting and the area under the peaks on ion chromatograms.[122,134] Spectral counting is the number of observed peptides normalised against the peptide in the reference sample that was most abundant providing a relative measure of protein quantity.[122] This means this method is semiquantative and can't account for sample loss.[135] Spectral counting is concentration specific and high abundance peptides will be preferentially measured while, low abundance proteins are usually lost or not significantly measured[121] Although low abundance peptides may be lost the reduction of proteins into peptides substantially increases the dynamic range.[121]

### Peptide and protein validation

Accurate identification of peptides and proteins is critical to the success of mass spectrometry analysis.[136] The large growth in mass spectrometry data sets has necessitated the development of complex algorithms to remove computational bottlenecks.[128] Multiple algorithms and software packages are available to efficiently process and convert raw binary based spectrum files such as .raw (Thermo Scientific), .mzML (HUPO-PSI) and .pkl (Waters) into usable text based mass spectrometric data such as MGF (Matrix Science) or custom XML files (X!Tandem and SEQUEST).[136,137] There are two critical steps involved in downstream processing: database searching and post processing.[136]

Peptide mass fingerprints (PMF) contained in the MS spectra and MS/MS spectra containing fragment ion $m/z$ ratios are generated and only contain peptide information such as peptide masses and relative abundances this must be converted back to protein identities by running the peptides against databases of *in silico* digested proteins created from gene information as shown in figure 1.8 below.[122,133,138] This increases the likelihood of identifying possible proteins but may increase the likelihood of false positives.[133] Specific parameters must be set to increase the likelihood of correct protein identification.[133] Parameters such as setting the taxonomy of the samples, what enzyme was used, missed cleavages from digestion and PTMs will reduce the likelihood of false positives.[133] This method also has the advantage that intrinsic information is retained, which can aid in identifying isoforms, polymorphisms and PTM's.[133] Carefully setting search parameters is critical for informative protein identification.[133]

Database search algorithms can be divided into three broad categories, fragment search/peptide mass fingerprinting, de novo search, and spectral library search algorithms.[128] Search engines such as Mascot, SEQUEST and X!Tandem match experimental spectra to theoretical proteins digested in silico.[128] The theoretical masses are then compared to the experimental masses to obtain peptide identities.[128]

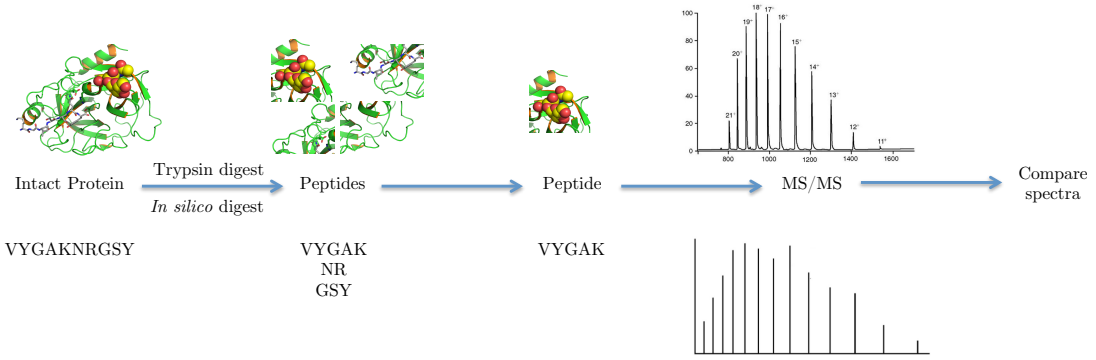


Figure 1.8: Schematic of the work flow of bottom up proteomics. Protein validation is achieved by digesting proteins using a chemical or enzymatic agents resulting in peptides, which are then, analysed using mass spectrometry. Theoretical mass spectra are also generated in silico and the experimental mass spectra is then compared theoretical mass spectra to obtain protein identities.

De novo searches such as NOVOR and PEAKS search through spectra previously identified and match the spectra to the best possible database entry.[128,139] This method can be extremely time consuming if the search space is not reduced within certain parameters.[128]

Spectral libraries such as MassBank use statistic parameters such as dot products or cross-correlations to reduce spectra to statistically defined statistical scores that are then searched and matched against a database.[128] The principle of spectral libraries dictates that protein fragmentation is statistically reproducible.[128,140]

Once the spectra has been matched and identified it must be processed to usable information for downstream analysis.[141] Software packages such as MaxQuant and PeptideShaker have been developed that can process spectrum files and their subsequent database data.[142,143] The focus of these packages is to assign significances and probabilities to the identification of the protein identities.[128] The use of multiple algorithms and software packages increases the effectiveness of database searching and post-processing.[136] Packages such as Scaffold and ProteoWizard have been developed that allow a more comprehensive understanding of mass spectrometric data.[144] Multiple new software packages are coming online each day and are making analysis easier and more effective.[145]

### Data filtering and quality control

Filtering of protein identification is critical in identifying significant matches.[136] Examples of filtering approaches include original-score based filtering, group, PeptideProphet, Percolator and local false discovery rate (LFDR). These are critical

in correct peptide matching.[136]

The most common quality control strategy in place is the target decoy strategy or original score based filtering whereby the spectrum data is shuffled into random combinations and run across a database. The subsequent proteins identified then act as a measure of how random the matching procedure is, resulting in the false discovery rate (FDR).[136] Further algorithms built on the FDR method were developed to increase the statistical significance of identified of peptide identities and reduce incorrect peptide spectrum matching (PSM).[136] PeptideProphet utilises a semi-supervised decoy method to estimate the probabilities of correct identities from discriminant scores coupled with a linear discriminant analysis based classifier.[136] Percolator uses a method whereby it identifies a subset of high confidence PSMs using a support-vector-machine based classifier to select correct PSMs.[136] LFDR utilises a log likelihood ratio generated using a naïve Bayesian classifier.[136]

### Experimental workflow

This study utilises a bottom up proteomic approach. An LC-ESI coupled to an Orbitrap Tribrid Fusion MS as depicted in figure 1.9 results in high throughput mass spectrometry approach.[121] This system utilises proteins digested to peptide fragments. These fragments were pre-separated by LC before being ionised using ESI. Initial MS1 was performed prior to selection of precursor ions which were then selected for fragmentation and fragmented using HCD. Fragment ions were subsequently analysed using the Orbitrap mass analyser. Detected fragmented ions were registered on binary .raw files which are and processed using the SequestHT algorithm to obtain peptide sequences and relative abundances were calculated and validated using X!Tandem The Orbitrap Tribrid Fusion architecture offers multiple fragmentation modes to produce product ions from precursor ions including collision-induced dissociation CID, higher-energy CID (HCD), electron-transfer dissociation and electron transfer HCD at any stage of MS.[126] Noise is actively reduced using an active beam guide by preventing neutral ions and high velocity ion clusters from entering the quadrupole mass filter.[126] MS/MS precursor ion selection can be achieved using the quadrupole mass filter with high ion transmission from 50 to 3000 $m/z$.[126] The Orbitrap mass analyser contains a nitrogen filled C-trap and improves resolving power up to 500 000 full width at half maximum (FWHM) and isotopic fidelity up to 240 000 FWHM at $m/z$ 200.[126] The ion-routing multipole allows HCD as well as ion storage without fragmentation and efficient and stable ion transfer between the ion routing multipole, the Orbitrap and the linear ion trap mass analysers which is critical component of Orbitrap mass analysers.[126] The dual pressure linear ion trap contains a high-pressure cell allowing MS precursor ion isolation from 0.2 amu to 600 amu and allows CID and ETD.[126] It also contains a low-pressure cell allowing improved scan speed, resolving

power and mass accuracy.[126] The dual dynode detector allows detection with a high linear dynamic range that allows improved quantitation.[126] This multiplexed architecture allows, an effective mass range of 50 - 6000 $m/z$., resolution from 15000 to 500 000 FWHM, MSn scans from 1 to 10 and a dynamic range greater than 5000 in a single scan.[126] The Orbitrap Tribrid Fusion MS in combination with a bottom-up approach has the distinct advantage of allowing the identification of protein identity and post-translational modifications of multiple proteins in the same run. By analysing complex samples from tumour and non-tumour samples a descriptive proteome can be identified that can reveal novel biomarkers for a disease state.[111]

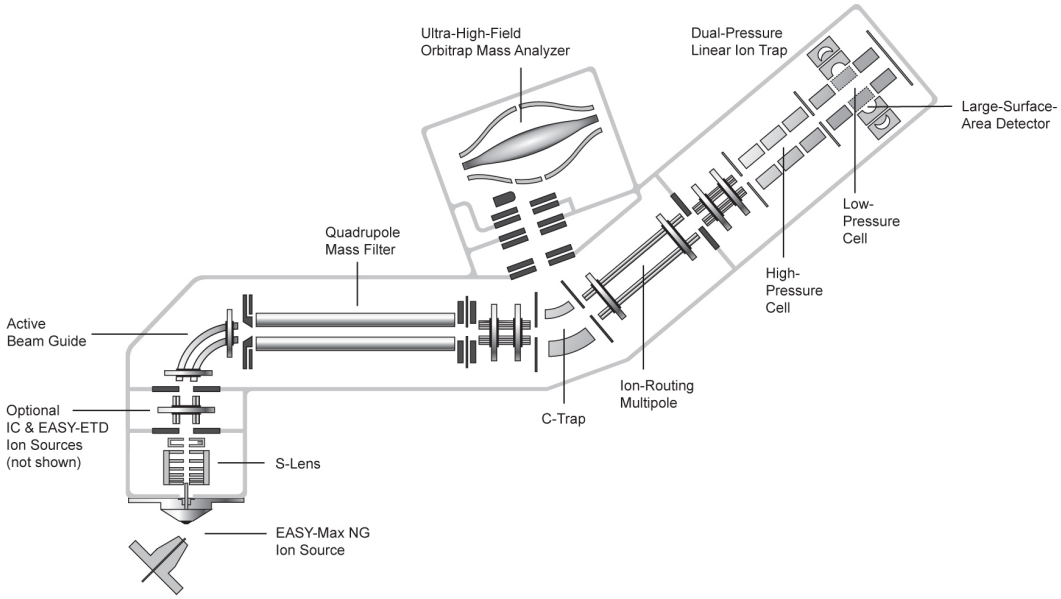


Figure 1.9: Schematic of an Orbitrap Fusion Tribrid mass spectrometer. The Tribrid architecture combines a quadrupole, dual pressure linear ion trap and an ultra-high-field Orbitrap mass analysers. Reproduced from Brooks/Cole, Cengage Learning.

## 1.5 Aims and objectives

The search for new non-invasive diagnostic tools for the identification of PCa continues. This study aims to identify possible candidate biomarkers for PCa through the elucidation of the proteome and secretome profiles of multiple prostate cell lines. The specific objectives of this study are:

1. To optimise the growth conditions of a normal prostate cell line (PNT2C2), a benign prostate hyperplasia cell line (BPH-1) and two prostate cancer cell lines (PC-3 and LNCaP) in order to increase the amount of secreted proteins produced while minimising cell death.

2. To optimise the detection rate of secreted proteins during mass spectrometry by optimising and selecting the most suitable protein precipitation method.

3. To characterise and compare the proteome and secretome of the four prostate cell lines with the aim of identifying changes unique to prostate cancer.

# Chapter 2

# Materials and methods

This study utilises an *in vitro* cell culture combined with mass spectrometry to identify the intracellular proteins (proteome) compared to the secreted proteins (secretome) with the goal of identifying possible candidate biomarkers to replace PSA to aid in the diagnosis of PCa.

## 2.1 General materials and reagents

Roswell Park Memorial Institute 1640 growth medium (RPMI-1640), glucose Hybrimax$^{TM}$, Ham's F12 Kaighn's modified growth medium (HAMSF12K) containing L-glutamine, acetone, methanol, chloroform, Coomassie brilliant blue G, phosphotungstic acid, NaCl, EDTA, guanidine-HCl, octylgluco-mano-pyranoside (OGP), methane methylthiosulphonate (MMTS), formic acid (FA) and trifluoroacetic acid (TFA) were all purchased from Sigma-Aldrich (St. Louis, USA). Sodium pyruvate, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and foetal bovine serum (FBS) were purchased from BioChrom (Cambridge, UK). Corning® 75 cm$^2$ and 175 cm$^2$ CellBIND® surface flasks and micro-centrifuge tubes were purchased from Corning® Life Sciences (NY, USA). T75 tissue culture flasks, 15 mL and 50 mL conical tubes were purchased from SPL Life Sciences (Gyeonggi-do, Korea). Bovine serum albumin (BSA) and ammonium sulphate were purchased from Roche Diagnostics (Mannheim, Germany) and Protea Chemicals (Western Cape, South Africa), respectively. Trypan blue stain, 1% Penicillin-Streptomycin (Pen-Strep), trypsin-ethylenediaminetetraacetic acid (trypsin-EDTA), chemically defined Chinese hamster ovary (CDCHO) medium and glutamine were purchased from Thermo-Fisher Scientific (Massachusetts, USA). A lactate dehydrogenase (LDH) standard was purchased from ScienCell Research Laboratories(California, USA). Triscarboxyethyl phosphine (TCEP), triethylammonium bicarbonate (TEAB), and UHPLC grade acetonitrile were purchased from Fluka (St. Louis, USA). Empore

C18 extraction discs were purchased from 3M(St. Louis, USA). Sequencing grade modified trypsin was obtained from Promega (Wisconsin USA).

## 2.2 Tissue culture

### 2.2.1 Cell line selection

Four cell lines representative of different states of the prostate were selected for the comparative analysis of their proteome and secretomes. The lymph node carcinoma of the prostate (LNCaP) representing androgen dependent malignant prostate cancer[146] and the human prostatic carcinoma (PC-3), representing androgen independent metastatic prostate cancer[147] cell lines were selected to differentiate between different stages of prostate cancer. The benign prostatic hyperplasia epithelial (BPH-1) cell line was selected to represent a benign condition often misdiagnosed as prostate cancer.[148] Normal prostate epithelium cells immortalised with SV40 (PNT2C2) were selected to represent healthy prostate cells.[149] LNCaP, PC-3 and PNT2C2 cell lines were purchased from the European Collection of Authenticated Cell Cultures (ECACC), while the BPH-1 cell line was a generous gift from S.W Hayward (Vanderbilt University, Nashville, Tennessee, USA).

### 2.2.2 Cell line maintenance

LNCaP cells were cultured in RPMI-1640 containing L-glutamine supplemented with 2.5 g/L D-(+)-Glucose, 20 mM HEPES and 2 mM sodium pyruvate. BPH-1 and PNT2C2 cells were cultured in RPMI-1640. PC-3 cells were cultured in HAMSF12K containing L-glutamine. Each growth medium was also supplemented with 10% FBS, 1.5 g/L sodium bicarbonate and 1% PenStrep. LNCaP cells were cultured in CellBIND® T75 tissue culture flasks, while all other cells were cultured in T75 tissue culture flasks. All cell lines were cultured at 37°C, 5% $CO^2$ and 90% humidity. Cell counts were determined using the Countess™ automated cell counter(Thermo-Scientific, California). All cell lines were routinely tested for mycoplasma infection using the indirect Hoechst stain technique and only mycoplasma negative cells were used for experiments.

### 2.2.3 Cell culture for proteome analysis

LNCaP, PC-3, BPH-1 and PNT2C2 cells were cultured for three passages to allow successful establishment of log phase cells and increase overall cell number. Five million cells of each cell line were subsequently seeded in 175$cm^2$ CellBIND® surface flasks and grown to 80% confluency (±5 days). The growth medium was

then aspirated and the cells collected using 4 mL trypsin-EDTA (0.05%) to detach the cells from the flasks. The cells were transferred to conical tubes containing 40 mL of the appropriate growth medium. The cells were centrifuged at 626 x g for 5 minutes and the growth medium aspirated and discarded. The cells were then resuspended in 20 mL of phosphate buffered saline (PBS, pH 7.4), vortexed and centrifuged at 626 x g for five minutes. The PBS was aspirated and discarded and the cells were resuspended in 1 mL of PBS and transferred to a 2 mL micro-centrifuge tube. The cells were then gently resuspended and centrifuged at 626 x g for five minutes. The PBS wash step was repeated in order to ensure that all traces of the growth medium was removed. The PBS was then aspirated and discarded and the cells weighed and aliquots of 50 mg each were stored at -80°C prior to mass spectrometry analysis. Three biologically independent samples were collected for each cell line.

### 2.2.4 Cell culture for secretome analysis

LNCaP, PC-3, BPH-1 and PNT2C2 cells were cultured for three passages as described above. Each cell line was subsequently seeded into three 175cm$^2$ CellBIND® surface flasks at three different seeding densities: LNCaP cells were seeded at 11,16.5 and 22 million cells per flask; PC-3 cells were seeded at 5, 8 and 11 million cells per flask; BPH-1 cells were seeded at 2, 5 and 8 million cells per flask; and PNT2C2 cells were seeded at 4, 6 and 8 million cells per flask. The cells were then incubated 48 hours, with the exception of the LNCaP cells, which were incubated for 72 hours. The regular growth medium was then aspirated and discarded. The cells were subsequently washed three times with 30 mL PBS in order to remove any traces of the original growth medium. CDCHO growth medium (30 mL) supplemented with 0.008 M glutamine was then added and the cells incubated for a further 48 hours. A clean T175 flask containing only CDCHO was also incubated under the same conditions to act as a negative control.

The CDCHO media was subsequently collected through gentle inversion of the 175cm$^2$ flask. The collected growth medium was centrifuged at 626 x g for five minutes to remove any excess cellular debris. The CDCHO media was subsequently transferred to into 50 mL conical tubes. The cells remaining in the flasks were collected using the aforementioned procedure, combined with the cellular debris and washed three times with 10 mL of PBS. The samples were snap frozen and stored at -80°C until later use. Three biologically independent samples were collected for each cell line.

## 2.3 Protein determination of secretome samples

Secretome samples were defrosted overnight at 4°C, vortexed and 1 mL aliquots collected. Protein concentration was determined using the Coomassie assay (Bradford).[150] Standards (0 $\mu$g/mL - 100 $\mu$g/mL) were prepared using BSA and MilliQ water ($mH_2O$). Standards and samples (20 $\mu$L) were added in triplicate to a 96-well flat-bottomed clear microtitre plate (Greiner Bio One International). Subsequently 100 $\mu$L of an in house Coomassie reagent was added. The plates were shaken on medium for three seconds (Biotek Power wave 340 microplate spectrophotometer) and incubated in the dark, at room temperature for 30 minutes. A modified approach was used to increase the sensitivity of the assay.[151] Absorbance was measured at 595 nm and 450 nm and a standard curve generated using the blank-corrected ratio of absorbance at 595 nm over 450 nm. The blank-corrected ratio of each sample at 595 nm and 450 nm was calculated and compared against the standard curve to calculate the protein concentration.

Table 2.1: Preparation of standards for the Bradford assay.

| Dilution | Volume of water added ($\mu$L) | Volume and source of BSA ($\mu$L) | Final BSA concentration ($\mu$g/mL) |
|---|---|---|---|
| 1 | 266 | 14 (of stock 2 mg/mL) | 100 |
| 2 | 140 | 140 (of dilution 1) | 50 |
| 3 | 140 | 140 (of dilution 2) | 25 |
| 4 | 140 | 140 (of dilution 3) | 12.5 |
| 5 | 140 | 140 (of dilution 4) | 6.25 |
| 6 | 140 | 140 (of dilution 5) | 2.125 |
| 7 | 140 | 140 (of dilution 6) | 1.5625 |
| 8 | 140 | 0 | 0 |
| 9 | 70 | 0 | 0 (No dye = Blank) |

## 2.4 Lactate Dehydrogenase determination

Cell death can be characterised by the quantification of plasma membrane damage. Lactate dehydrogenase is a stable cytoplasmic enzyme that is released from the cytosol of cells into the surrounding environment during cell death. Its concentration is therefore a useful measurement of cell death and cytotoxicity. LDH levels can be

measured using a colorimetric assay. During the assay reduced nicotinamide adenine dinucleotide ($NAD^+$) is reduced to oxidised nicotinamide adenine dinucleotide ($NADH + H^+$) by the LDH-catalysed conversion of lactate to pyruvate. In the presence of the catalyst (diaphorase) $NADH + H^+$ in turn reduces tetrazolium salt (2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyltetrazolium chloride) (INT) to form a formazan dye product. An increase in the amount of plasma membrane damaged cells will cause an increase in the LDH enzyme which correlates to an increase in the formazan dye product which can be detected at 495 nm using a microplate spectrophotometer.[152] LDH concentration was determined using the lactate dehydrogenase assay from the cytotoxicity detection kit (Roche). Standards (0 mU/mL - 200 mU/mL) were prepared using a LDH standard diluted with CDCHO media from the negative control samples described above. Each standard (10 $\mu$L) and sample (10 $\mu$L) were added in triplicate to a 96 well microtitre plate followed by the addition of 100 $\mu$L of the reaction mixture to each well. Samples were shaken on medium for three seconds and incubated in the dark for 30 minutes at room temperature. Absorbance was subsequently measured at 495 nm and a standard curve prepared. The relative levels of LDH in the samples were determined by comparing the blank-corrected absorbance of the samples to the blank-corrected standard curve. Samples with absorbance values, which were higher than the highest value on the standard curve, were diluted using $mH_2O$ and the assay was repeated.

Table 2.2: Preparation of standards for the LDH assay.

| Dilution | Volume of CDCHO added ($\mu$L) | Volume and source of LDH ($\mu$L) | Final LDH concentration ($\mu$g/mL) |
|---|---|---|---|
| 1 | 560 | 140 (of stock 1 U/mL) | 200 |
| 2 | 350 | 350 (of dilution 1) | 100 |
| 3 | 350 | 350 (of dilution 2) | 50 |
| 4 | 350 | 350 (of dilution 3) | 25 |
| 5 | 350 | 350 (of dilution 4) | 12.5 |
| 6 | 350 | 350 (of dilution 5) | 6.25 |
| 7 | 350 | 350 (of dilution 6) | 3.125 |
| 8 | 350 | 0 | 0 = Blank |

## 2.5 Protein precipitation of secretome samples

The proteins in the secretome samples were precipitated in order to increase the detection of low abundance proteins during mass spectrometry analysis. Four methods of protein precipitation were investigated using pooled secretome samples. These included precipitation using acetone, ammonium sulphate, methanol/chloroform and phosphotungstic acid (PTA) mediated acetone precipitation.

### 2.5.1 Acetone precipitation

Acetone was cooled overnight to -20°C. The secretome samples (27 mL) were transferred to 250 mL glass bottles and 108 mL of four volumes of acetone were added to each sample. The samples were mixed and incubated for 24 hours at -20°C. The samples were subsequently centrifuged in 40 mL aliquots at 15000 xG for 10 minutes and the supernatant was removed and discarded. The remaining protein pellet was allowed to air-dry and was then resuspended in 1 mL $mH_2O$. The overnight reaction vessel was washed with 5 mL $mH_2O$ and transferred to the pelleted sample to collect any remaining sample. The samples were then transferred to 15 mL conical tubes and the original centrifuge tubes were washed with 1 mL $mH_2O$ in order to collect any remaining sample. The samples were dried overnight using a speedy-vac (LabConco Centrivap) to near dryness and then resuspended in 2 mL $mH_2O$. Protein concentration was determined using the aforementioned procedure and samples snap frozen and stored at -80°C until later use.

### 2.5.2 Ammonium sulphate precipitation

Samples (27 mL) were transferred to centrifuge 50 mL centrifuge tubes. Ammonium sulphate (14.39 g) was added over a period of an hour while mixing in order to achieve 80% saturation. The samples were then incubated for an hour at 4°C to allow for protein precipitation. The samples were subsequently centrifuged at 20000 xG for 30 minutes at 4°C and the supernatant was removed and discarded. The protein pellets were resuspended in 5 mL $mH_2O$ and transferred to 15 mL conical tubes. The conical tube lids were pierced with a 0.5 cm radius hole punch. Dialysis tubing (2500 Da) (Thermo-Scientific Snakeskin CA) was captured between the lid and conical tube and checked for leaks through slow inversion. The dialysis vessels were placed in 500 mL $mH_2O$, 4°C and allowed to dialyse overnight at 4°C under constant light magnetic stirring. The $mH_2O$ was then changed and the samples were dialysed for a further 4 hours after which the $mH_2O$ was again changed and the samples dialysed for a further 4 hours. The samples were subsequently freeze-dried to near dryness overnight using a speedy-vac, resuspended in 2 mL $mH_2O$ and transferred to 2 mL micro-centrifuge tubes. Protein concentration was determined

using the aforementioned procedure and samples snap frozen and stored at -80°C until later use.

### 2.5.3 Methanol chloroform precipitation

Samples were transferred to 500 mL centrifuge bottles. Four volumes of methanol (108 mL) were added to the samples and samples were mixed vigorously. One volume of chloroform (27 mL) was then added to each sample. Three volumes of $mH_2O$ (81 mL) were subsequently added to the samples, which were then vortexed. Samples were subsequently centrifuged at 14000 xG for 2 minutes. The top aqueous layer was aspirated and discarded. Four volumes of methanol (108 mL) were added to the samples and vortexed. The samples were centrifuged at 14000 xG for 3 minutes and the remaining methanol was aspirated and discarded. The remaining sample was dissolved in 5 mL $mH_2O$, transferred to 15 mL conical tubes, dried overnight to near dryness and then resuspended in 2 mL $mH_2O$ and transferred to 2 mL micro-centrifuge tubes. Protein concentrations were determined using the aforementioned procedure and samples snap frozen and stored at -80°C until later use.

### 2.5.4 Phosphotungstic acid (PTA) mediated acetone precipitation

The procedure was identical to acetone precipitation except that 1.2% phosphotungstic acid (w/v) was added to the acetone sample reaction vessel prior to overnight incubation. Protein concentration was determined using the aforementioned procedure and samples snap frozen and stored at -80°C until later use.

## 2.6 Mass spectrometric analysis

The 50mg whole cell and acetone-precipitated samples were sent to the Stellenbosch Central Analytical Facility, Tygerberg, South Africa for mass spectrometric analysis.

### 2.6.1 Whole cell sample preparation

Cells were thawed in 100 $\mu$L extraction buffer and proteins extracted by grinding the pellets with glass beads in extraction buffer containing; 100 mM NaCl, 2 mM EDTA, 6 M guanidine-HCl, 1% (v/v) OGP and 5 mM TCEP in 100 mM TEAB (pH 7.8). After removal of the remaining pellet was washed with 100 $\mu$L of 100 mM TEAB. The supernatants from were pooled and an overnight acetone precipitation was performed with 5 volumes acetone. Thereafter, the samples

underwent centrifugation the pellets were air dried and dissolved in 100 $\mu$L 100 mM TEAB containing 4 M guanidine-HCl and 1% OGP (v/v). Protein concentrations were determined spectrophotometrically.

### 2.6.2 Sample digests

Samples were reduced by adding 1 $\mu$L of 50 mM TCEP in 90 $\mu$L 100 mM TEAB (final concentration 5 mM TCEP) for 30 minutes at room temperature. Following reduction cysteine residues were modified to methylthio using 1 $\mu$L 200 mM MMTS in 10 $\mu$L 100mM TEAB (final concentration 10 mM) for 30 minutes. After modification, the samples were diluted to 98 $\mu$L with 100 mM TEAB. Proteins were digested by adding 5$\mu$L 1 $\mu$g/$\mu$L trypsin solution (1:20 ratio) and incubating for 18 hours at 37°C. The samples were dried down and resuspended in 100 $\mu$L 2% (v/v) acetonitrile:water; 0.1% (v/v) FA.

### 2.6.3 Sample desalting

Residual digest reagents were removed using an in-house manufactured C18 stage tip using C18 extraction discs The 20 $\mu$L sample was loaded onto the stage tip after activating the C18 membrane with 30 $\mu$L methanol and equilibration with 30 $\mu$L 2% (v/v) acetonitrile:water; 0.1 FA. The bound sample was washed with 30 $\mu$L 2% (v/v) acetonitrile:water; 0.1% (v/v) FA before elution with 30 $\mu$L 50% (v/v) acetonitrile:water 0.1% (v/v) FA. The eluate was evaporated to dryness. The dried peptides were dissolved in 20 $\mu$L 2% (v/v) acetonitrile:water; 0.1% (v/v) FA for liquid chromatography mass spectrometry (LC-MS) analysis.

### 2.6.4 Liquid chromatography

Liquid chromatography was performed on a Thermo Scientific Ultimate 3000 RSLC equipped with a 0.5 cm x 300 $\mu$m ID C18 trap column and a 40 cm x 75 $\mu$m ID in-house manufactured C18 column (Luna C18, 3.6 $\mu$M; Phenomenex) analytical column. The solvent system employed was loading: 2% (v/v) acetonitrile:water; 0.1% (v/v) FA; Solvent A: 2% (v/v) acetonitrile:water; 0.1% (v/v) FA and Solvent B: 100% (v/v) acetonitrile:water; 0.1% (v/v) FA. The samples were loaded onto the trap column using loading solvent at a flow rate of 15 $\mu$L/minute from a temperature controlled autosampler set at 7°C. Loading was performed for 5 minutes before the sample was eluted onto the analytical column. Flow rate was set to 500 nL/minute and the gradient generated as follows: 2.0%-10% B over 5 min; 5%-25% B from 5-50 minutes using Chromeleon non-linear gradient 6, 25%-45% from 50-65 minutes, using Chromeleon non-linear gradient 6. Chromatography was performed

at 50°C and the outflow delivered to the mass spectrometer through a stainless steel nano-bore emitter.

### 2.6.5 Mass spectrometry of whole cell samples

Mass spectrometry was performed using a Thermo Scientific Fusion mass spectrometer equipped with a Nanospray Flex ionisation source. The sample was introduced through a stainless steel emitter. Data was collected in positive mode with spray voltage set to 2 kV and ion transfer capillary set to 275°C. Spectra were internally calibrated using polysiloxane ions at $m/z = 445.12003$ and 371.10024. Mass spectrometry scans were performed using the Orbitrap detector set at 120 000 resolution over the scan range 350-1650 with AGC target at $3x10^5$ and maximum injection time of 40 ms. Data was acquired in profile mode.

Tandem mass spectrometry acquisitions were performed using monoisotopic precursor selection for ion with charges +2 - +6 with error tolerance set to ±10 ppm. Precursor ions were excluded from fragmentation once for a period of 30 seconds. Precursor ions were selected for fragmentation in higher-energy collisional dissociation mode using the quadrupole mass analyser with HCD energy set to 32.5%. Fragment ions were detected in the Orbitrap mass analyser set to 15 000 resolution. The automatic gain control target was set to $1x10^4$ and the maximum injection time to 45 ms. Data was acquired in centroid mode.

### 2.6.6 Data analysis

The raw files generated by the mass spectrometer were imported into Proteome Discoverer v1.4 (Thermo-Scientific, California) and processed using the SequestHT (Thermo-Scientific, California) algorithm included in Proteome Discoverer. Settings allowed for methylthio as fixed modification as well as NQ deamidation (NQ) and oxidation (M). Precursor tolerance was set to 10 ppm and fragment ion tolerance to 0.02 Da. The Homo sapiens database from the Universal protein resource (UniProt) was used for analysis. The results files were imported into Scaffold 1.4.4 (Proteome Software, Oregon) and identified peptides validated using the X!Tandem search algorithm included in Scaffold.

For comparative analysis between cell lines to ensure correct identification of the proteome profile Stringent search engine constraints were set for the Sequest database (Thermo-Scientific, California) to ensure correct identification of the proteome profile. The fragment tolerance was set at 0.020 Da (monoisotopic), precursor tolerance at 10.0 ppm (monoisotopic), fixed modifications at +46 on C (methylthio), variable modification at +1 on NQ (deamidated), +16 on M (oxidation) and +42 on n (acetyl) and maximum missed cleavages were set at 2.

X! Tandem database identified proteins using the PeptideProphet algorithm (The GM, thegpm.org) with Scaffold delta-mass correction with fragment tolerance set at 0.020 Da (monoisotopic), precursor tolerance at 10.0 ppm (monoisotopic), fixed modifications at +46 on C (methylthio), variable modification at -18 on n (Glu->puro-Glu), -17 on n (Gln->pyro-Glu) and +1 on NQ (deamidated), +16 on M (oxidation), +42 on n (acetyl) maximum missed cleavages at 2.

Peptide probabilities were assigned by the Scaffold Local FDR by Scaffold (Proteome Software, Oregon). Constraints were set at 95% minimum for peptide thresholds, 99% minimum for protein threshold and a minimum of 2 peptides, 0.0% peptide FDR and 0.5% for protein FDR.

Proteins identifications were only accepted if the probability was above 99% and contained at least 2 identified peptides. Proteins the contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony and proteins sharing significant peptide evidence were grouped into clusters and later excluded prior to analysis.

Peptide and Protein validation were done using the Peptide and ProteinProphet algorithms. GO annotations were retrieved from the NCBI database.

### 2.6.7 Ingenuity® Pathway Analysis

The proteins identities for BPH-1, LNCaP and PC-3, their relative $\log_2$ fold change ratios compared to PNT2C2 and their significance measured using the Benjamini-Hochberg multiple test correction with a significance of $p < 0.05$ were analysed using QIAGEN's Ingenuity® Pathway Analysis (IPA® QIAGEN Redwood City, www.qiagen.com/ingenuity)(IPA). A core analysis was run independently on each cell line. Analysis parameters were set as follows: the reference set was the Ingenuity Knowledge Base analysing direct and indirect relationships, interaction networks were set to include endogenous chemicals with 35 molecules per network and 25 networks per analysis, all node types were measured, all data sources were included, confidence was set at experimentally observed, species were set for all including: human, mouse, rat and uncategorised, tissues and cell lines were ignored and mutations were set to all. Focus was done on both upregulated and downregulated molecules and experimental significance was set at $p < 0.05$. Canonical pathways, upstream analysis, diseases and functions, regulator effects and networks were reviewed and significant data was selected.

Comparative analyses were performed on the BPH-1, LNCaP and PC-3 cell lines. Comparative canonical pathways, upstream analysis, diseases and functions, regulator effects and networks were all reviewed. Biomarker analysis was performed on proteins identified as unique to the LNCaP and/or PC-3 cell lines.

# Chapter 3

# Results

## 3.1 Cell culture optimisation

Previous studies have identified working seeding densities for the LNCaP and PC-3 cell lines and revealed that lower seeding densities generally result in lower levels of cell death with comparatively higher levels of protein secretion.[153] This study builds on these findings by confirming the optimal seeding densities for the LNCaP and PC-3 and investigating working seeding densities for the BPH-1 and PNT2C2 cell lines.

Optimal seeding densities were therefore confirmed prior to protein precipitation and subsequent mass spectrometric analysis. The BPH-1, LNCaP, PC-3 and PNT2C2 cell lines were each seeded at three different densities. After three rounds of sub-culturing the normal media was removed, the cells were washed with PBS and media was replaced with CDCHO media to minimise the presence proteins not produced by the cells. After 48 hours (72 hours in the case of LNCaP cells) the CDCHO media was collected and protein and LDH concentrations were determined to ascertain which seeding density had the least amount of cell death with the largest amount of protein secretion.

As might be expected the results indicate that protein concentration was highest in the highest seeding densities with the exception of the difference between the $16.5x10^6$ and $22x10^6$ seeding density for LNCaP cell which was not significant as shown in figure 3.1. LDH concentration was the lowest in the lowest seeding densities for all cell lines with the exception of the LNCaP cells with showed no significant difference between the seeding densities as shown in figure 3.2. The ratio of LDH to protein concentration (figure 3.3) revealed that the optimal seeding densities, as indicated by the lowest ratio of LDH to protein concentration, would be the lowest seeding density for the BPH-1 and PC-3 cell lines. While the lowest ratios tended to be obtained for the highest seeding densities for both the LNCaP

and PNT2C2 cell lines these seeding densities were not significantly different from the other seeding densities.

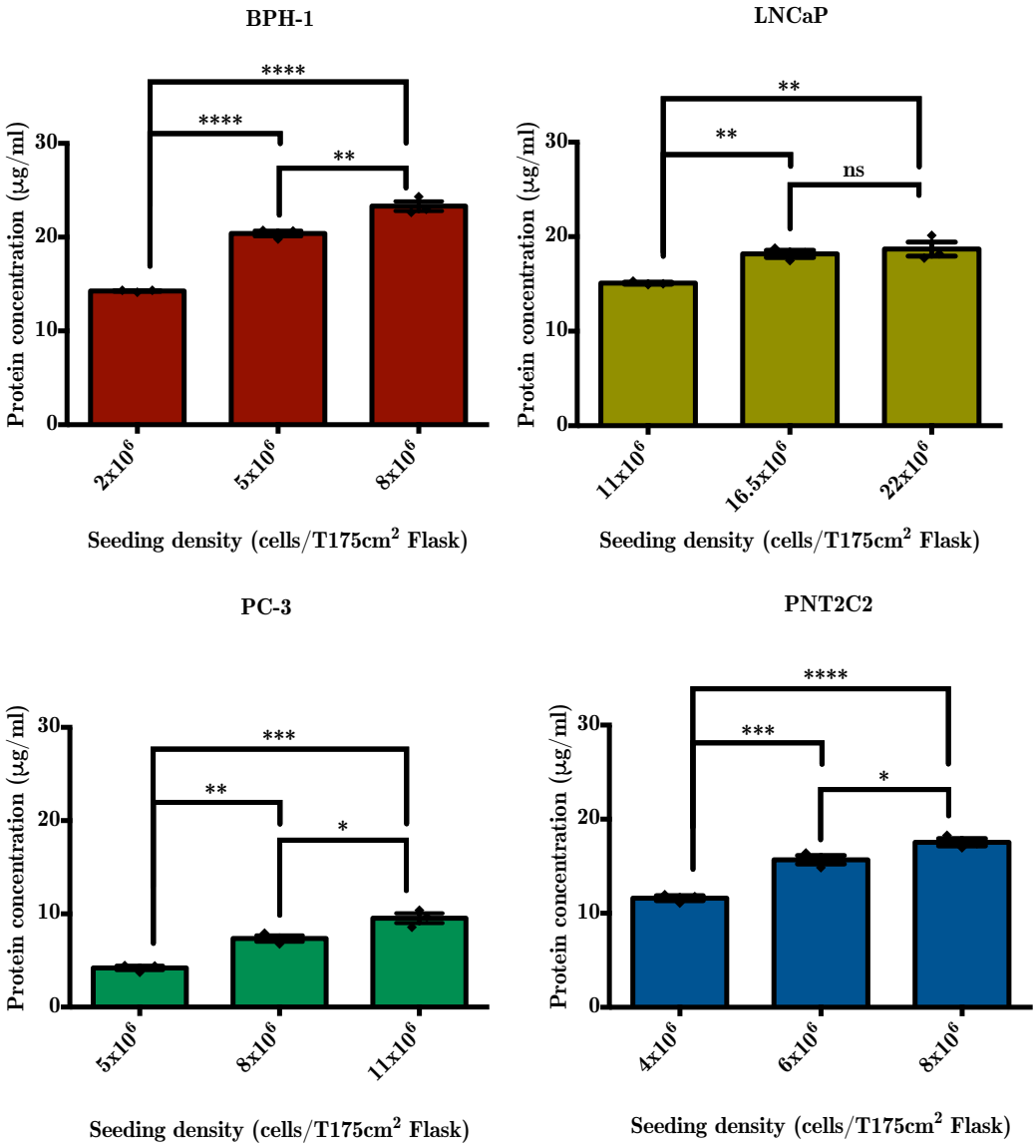


Figure 3.1: Protein concentrations in CDCHO media conditioned using multiple seeding densities of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines. Results are shown as the means ± SEM of three independent experiments each measured in triplicate. Tukey's post hoc test was applied and significance is indicated as ns ($p > 0.05$), * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) and **** ($p < 0.0001$).

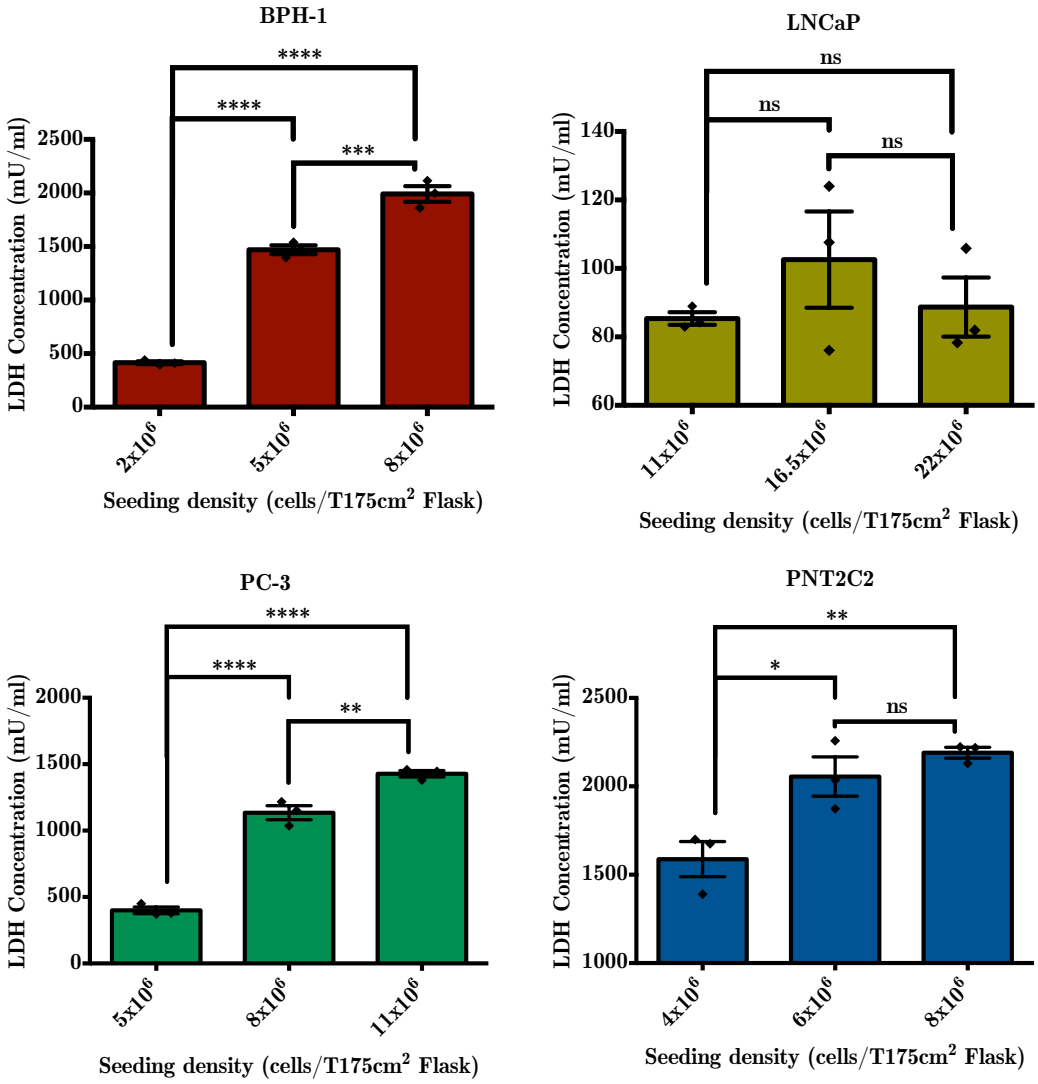


Figure 3.2: LDH concentrations in CDCHO media conditioned using multiple seeding densities of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines. Results are shown as the means ± SEM of three independent replicates measured in triplicate. Tukey's post hoc test was applied and significance is indicated as ns ($p > 0.05$), * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) and **** ($p < 0.0001$).

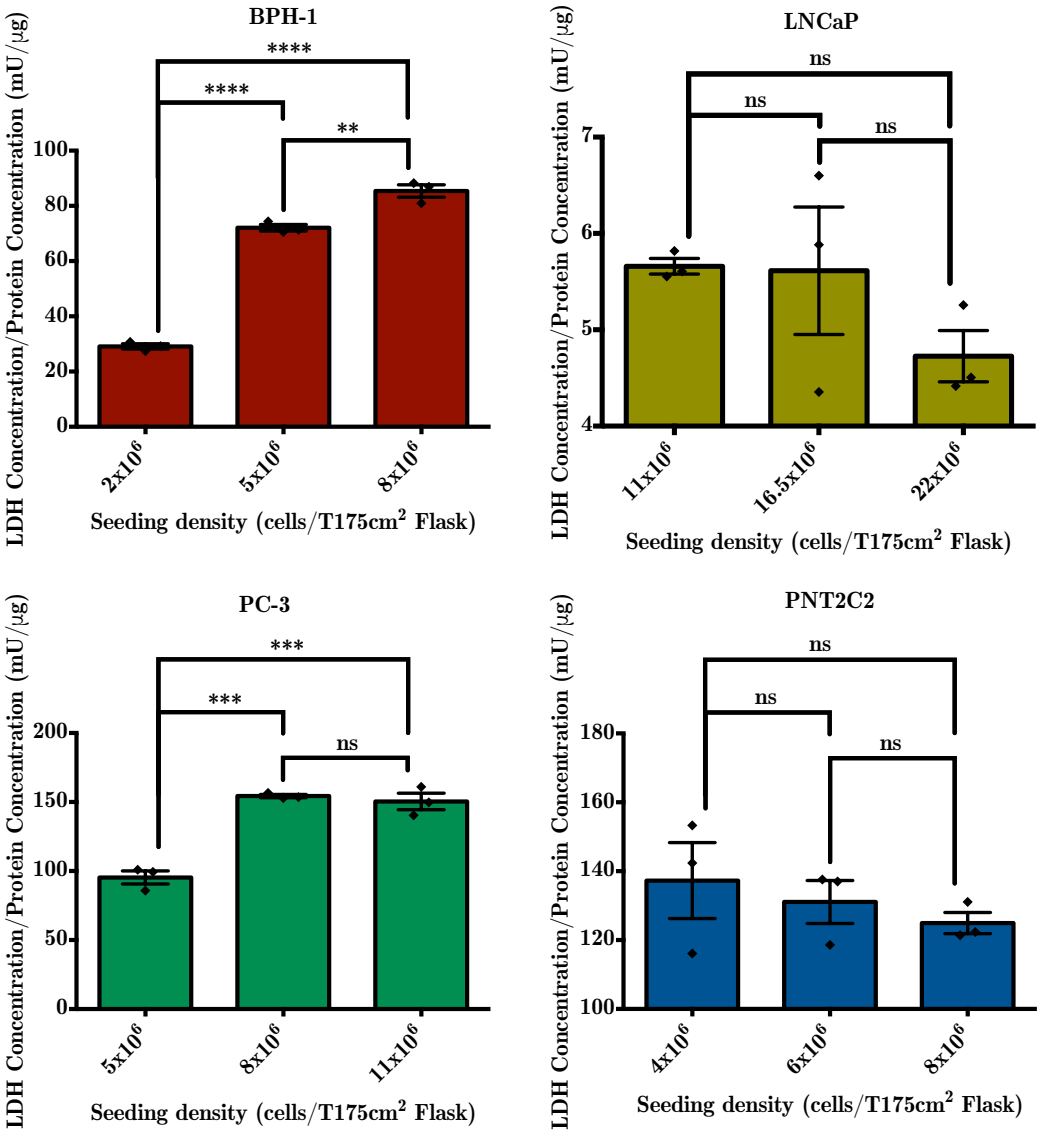


Figure 3.3: LDH/Protein concentrations ratios in CDCHO media conditioned using multiple seeding densities of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines. Results are shown as the means ± SEM of three independent replicates measured in triplicate. Tukey's post hoc test was applied and significance is indicated as ns ($p > 0.05$), * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) and **** ($p < 0.0001$).

## 3.2 Protein precipitation

Protein precipitation methods are routinely used to increase the concentrations of proteins, reduce the amount of solvent, dissociate proteins from carrier molecules and to aid in the removal of contaminants.[154,155]

Four methods of precipitation were tested in this study in order to identify the most appropriate method. These included: acetone precipitation; ammonium sulphate precipitation; methanol/chloroform precipitation and PTA mediated acetone precipitation. Samples from each seeding density were pooled for each cell line and subsequently separated into five aliquots; four test samples and one control. The resulting aliquots from each of the four cell lines were then used to compare the four precipitation methods.

The results shown in table 3.1 and figure 3.4 indicate that acetone precipitation method yielded the highest recovery while ammonium sulphate precipitation resulted in the lowest measurable recovery. No protein was recovered from methanol/chloroform precipitation as the methodology employed prevented the collection of the protein interphase, while the protein concentration from the PTA mediated acetone precipitation could not be accurately determined and this method was thus excluded. Figure 3.4 lists the protein recovery as a percentage of protein recovered compared to the control sample.

Table 3.1: Comparison of three methods used to precipitation proteins from conditioned CDCHO medium. Samples were created by pooling three secretome seeding densities from each cell line. The control was an aliquot of the pooled samples that did not undergo precipitation. Results are shown as the means ± SEM of four independent replicates measured in triplicate. Dunnetts' post hoc test was applied and significance is indicated as ns ($p > 0.05$), * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) and **** ($p < 0.0001$).

| | Total recovery ($\mu$g/ml) | | |
|---|---|---|---|
| Control ($\mu$g/ml) | Ammonium Sulphate | Acetone | Methanol + Chloroform |
| 286 ± 66.4 | 90.9 ± 7.85 | 161 ± 10.7 | 0 ± 0 |
| Significance | ** | ns | *** |

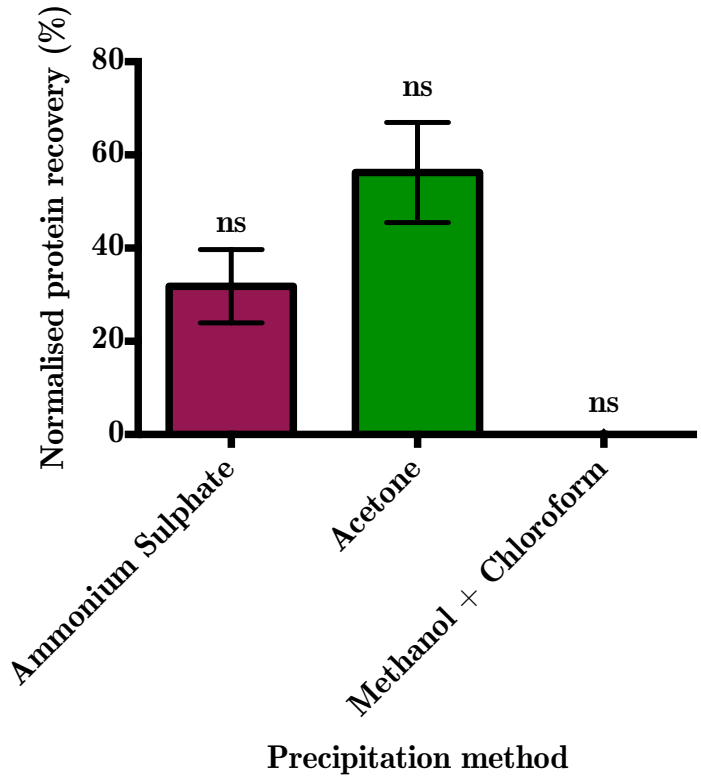


Figure 3.4: The percentage recovery obtained using three methods to precipitate proteins from conditioned CDCHO medium. Results were normalised against the appropriate control sample by setting the control to 100%. Results are shown as the means ± SEM of four independent experiments each measured in triplicate. Dunnetts' post hoc test was applied and significance is indicated as ns ($p > 0.05$), * ( $p < 0.05$), ** ( $p < 0.01$), *** ($p < 0.001$) and **** ($p < 0.0001$).

## 3.3 Proteome analysis

### 3.3.1 Mass spectrometry guidelines

This study builds on countless previous works with the aim of fully categorising tissue culture cell lines to aid in biomarker development as understand functional differences between the cell lines.[153,156] Furthermore, the analysis of cell line secretomes is of great value for biomarker discovery as it is simpler than the analysis of clinical samples which are complicated by the presence of high abundance proteins which can mask the measurement of lower abundance proteins.[157]

While care was taken to minimise the amount of lysed cells during the collection

of secretome samples in this study, cell lysis is unavoidable. The approach used in this study was therefore to characterise the proteome as well as the secretome and in so doing eliminate false positive biomarkers that may have contaminated the secretome proteomic data. The analysis of the proteome also allows for differences in functional pathways to be compared between the cell lines.

The BPH-1, LNCaP, PC-3 and PNT2C2 cell lines were all grown under optimal conditions. After four rounds of sub-culturing the cells the cells were collected and washed and aliquots prepared for mass spectrometric analysis. Three independent biological replicates for each cell line were each injected in triplicate. The resulting mass spectra were subsequently analyses using the SequestHT and X!Tandem search engines and the integrated algorithms available in Scaffold and PeptideProphet using stringent search parameters. All the data was subsequently imported into Scaffold and protein identities were assigned using stringent parameters. Thresholds were set at 95% minimum for peptide and 99% minimum for proteins. Proteins were assigned using a minimum of 2 peptides and a FDR of 0.5%. The data acquired in this study meets the standards The Minimum Information About a Proteomic Experiment (MIAPE) guidelines[158–160]

The conversion rate from spectra to ID means were 18%, 20%, 21% and 19% for BPH-1, LNCaP, PC-3 and PNT2C2 respectively.

### 3.3.2 Comparison of protein expression between cell lines

A total of 3576 proteins were identified in this study. Of these 1253 (35%) were common to all cell lines, while 216 (6%), 431 (12%), 284 (8%) and 238 (7%) proteins were unique to the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines, respectively. The overlap of identified proteins is shown below in the Venn diagram (figure 3.5).

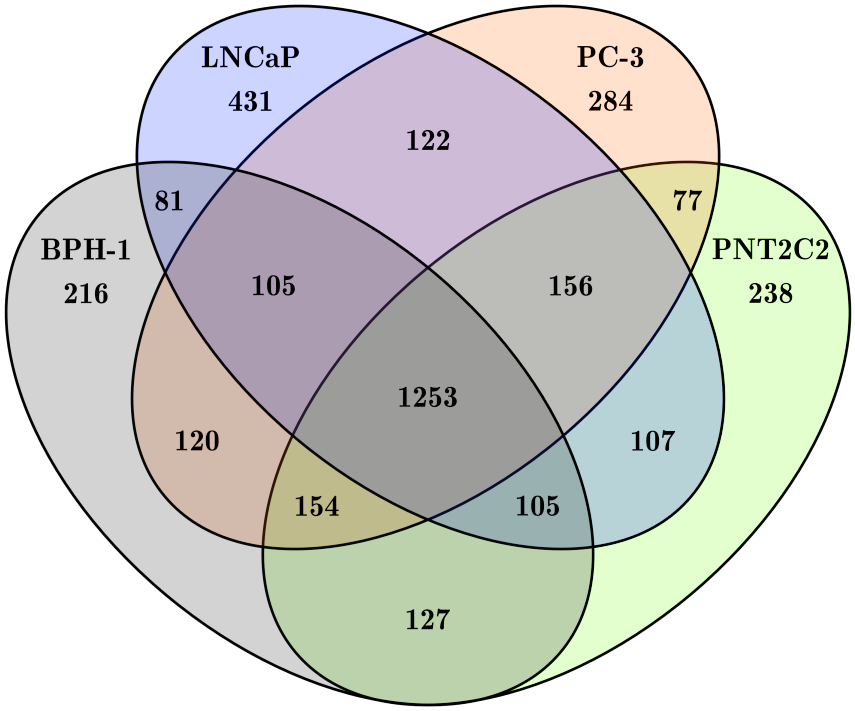


Figure 3.5: Venn diagram showing the overlap of all proteins identified in the proteome of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines.

Scaffold was subsequently used to quantify the expression levels of proteins (log2 fold change changes) in the BPH-1 (benign prostate hyperplasia), LNCaP (androgen dependent prostate cancer) and PC-3 (androgen independent prostate cancer) cell lines relative to the reference cell line PNT2C2 (normal prostate). A T-Test with a Benjamini-Hochberg multiple test correction (n=36) was performed by scaffold and p-values were determined for the fold change of each protein. This data was subsequently analysed using the IPA software package.

### 3.3.3 IPA data analysis

Protein information was uploaded to IPA and 92% (BPH-1), 93% (LNCaP) and 93% (PC-3) of proteins were correctly mapped to annotated identities from the GenPept and UniProt/Swiss-Prot accession databases. All unidentified, non-annotated and duplicated identities were discarded. Protein expression of the BPH-1, LNCaP and PC-3 cell lines were compared against PNT2C2 protein expression and measured using the $\log_2$ fold change ratio and a p-value cut off of 0.05 for all analyses. A total of 51%, 52% and 53% of the proteins identified in the BPH-1, LNCaP and

PC-3 cell lines were found to be expresses at a level which was significantly different ($p<0.05$) to that of the PNT2C2 cell line.

### Canonical pathway analysis

Canonical pathways represent idealised or generalised pathways that represent common properties of a particular signalling module or pathway.[161] Canonical pathways give an indication of the changes of cellular functions and biological pathways that can give insight into the pathways responsible for observed differences in protein expression between cell lines. Tables 3.2, 3.3 and 3.4 list the top five canonical pathways identified in the BPH-1, LNCaP and PC-3 cell lines, respectively which were significantly different when compared to the PNT2C2 reference cell line.

The results show that for the BPH-1 cell line the EIF2 signalling pathway, the regulation of the eIF4 and p70S6K signalling, the protein ubiquitination pathway, the mTOR pathway and the tRNA charging pathways were the most changed with p-values ranging from $1.68x10^{-08}$ to $9.77x10^{-39}$and pathway coverage ranging from 19.1% to 33.3% (Table 3.2).

Table 3.2: Top five canonical pathways altered in the BPH-1 cell line relative to the PNT2C2 cell line.

| Name | p-value | Overlap (%) | Overlap |
|---|---|---|---|
| EIF2 signalling | $9.77x10^{-39}$ | 34.4 | 76/221 |
| Regulation of eIF4 and p70S6K signalling | $2.30x10^{-22}$ | 30.6 | 48/157 |
| Protein ubiquitination pathway | $1.10x10^{-15}$ | 20.4 | 52/255 |
| mTOR signalling | $8.26x10^{-13}$ | 20.6 | 41/199 |
| tRNA charging | $1.68x10^{-08}$ | 35.9 | 14/39 |

The EIF2 signalling pathway, the mitochondrial dysfunction pathway, oxidative phosphorylation proteins, the eIF4 and p70S6K signalling pathways and the remodelling of epithelial adherens junctions pathway were the the most changed in the LNCaP cell line with p-values ranging from $2.66x10^{-11}$ to $4.93x10^{-28}$ and pathway coverage ranging from 26.9% to 32.4%. (Table 3.3).

Table 3.3: Top five canonical pathways altered in the LNCaP cell line relative to the PNT2C2 cell line.

| Name | p-value | Overlap (%) | Overlap |
|---|---|---|---|
| EIF2 signalling | $4.93x10^{-28}$ | 25.8 | 65/221 |
| Mitochondrial dysfunction | $1.72x10^{-18}$ | 25.1 | 46/171 |
| Oxidative phosphorylation | $2.68x10^{-15}$ | 27.5 | 33/109 |
| Regulation of eIF4 and p70S6K signalling | $8.17x10^{-12}$ | 15.7 | 35/157 |
| Remodelling of epithelial adherens junctions | $2.66x10^{-11}$ | 19.1 | 22/68 |

The top five regulated pathways in the PC-3 cell line were the protein ubiquitination pathway, the EIF2 signalling pathway, the remodelling of epithelial adherens junctions, tRNA charging and the epithelial adherens junction signalling pathways, with p-values ranging from $7.66x10^{-11}$ to $1.18x10^{-16}$ and pathway coverage ranging from 21.2% to 43.6%. (Table 3.4).

Table 3.4: Top five canonical pathways altered in the PC-3 cell line relative to the PNT2C2 cell line.

| Name | p-value | Overlap (%) | Overlap |
|---|---|---|---|
| Protein ubiquitination pathway | $1.18x10^{-16}$ | 21.2 | 54/255 |
| EIF2 signalling | $4.47x10^{-16}$ | 22.2 | 49/221 |
| Remodelling of epithelial adherens junctions | $2.99x10^{-13}$ | 35.3 | 24/68 |
| tRNA charging | $1.64x10^{-11}$ | 43.6 | 17/39 |
| Epithelial adherens junction signalling | $7.66x10^{-11}$ | 21.9 | 32/146 |

The relative proportions of proteins that are upregulated, unchanged and downregulated in the BPH-1, LNCaP and PC-3 cell lines when compared to PNT2C2 are shown in figures 3.6, 3.7 and 3.8, respectively.

This data shows that in the BPH-1 cell line 42%, 35%, 18%, 26% and 46% of proteins matched to the EIF2 signalling pathway, the eIF4 and p70S6K signalling pathways, the protein ubiquitination pathway, the mTOR signalling pathway and the tRNA charging pathway, respectively were downregulated when compared to PNT2C2, while 14%, 13%, 16%, 10% and 8% of the proteins were upregulated. (Figure 3.6).

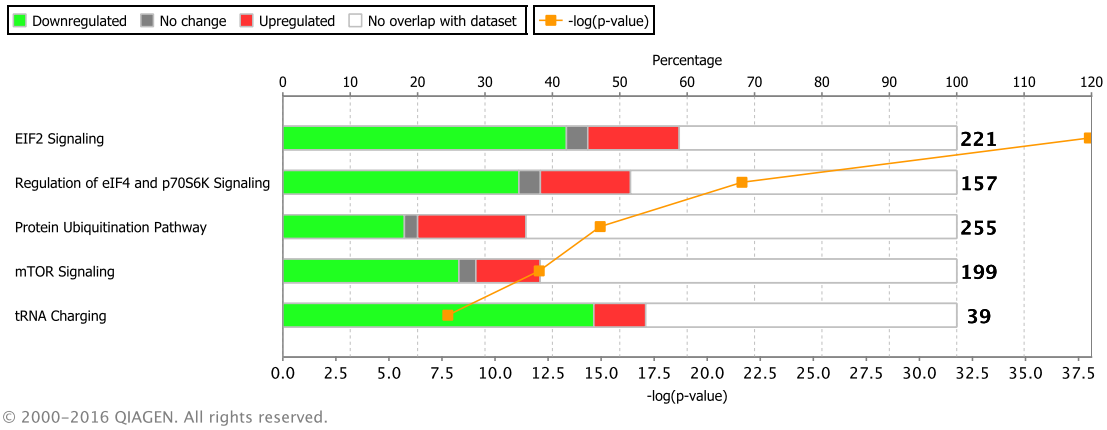




Figure 3.6: Top five canonical pathways altered in the BPH-1 cell line relative to the PNT2C2 cell line. Green segments, grey segments and red segments indicate the percentage of proteins that are downregulated, unchanged, or upregulated, respectively when compared to the PNT2C2 cell line. The -log(p-value) indicates the significance of the changes of BPH-1 when compared to PNT2C2 and was calculated as a right tailed Fisher's exact. Values at the end of the bars indicate the total number of proteins in the respective pathways.

In the LNCaP cell line 16%, 2%, 4%, 15% and 41% of proteins matched to the EIF2 signalling pathway, the mitochondrial dysfunction pathway, the oxidative phosphorylation pathway, the protein ubiquitination pathway and the eIF4 and p70S6K signalling pathways, respectively were downregulated when compared to PNT2C2, while 33%, 36%, 39%, 28% and 13% of the proteins were upregulated.

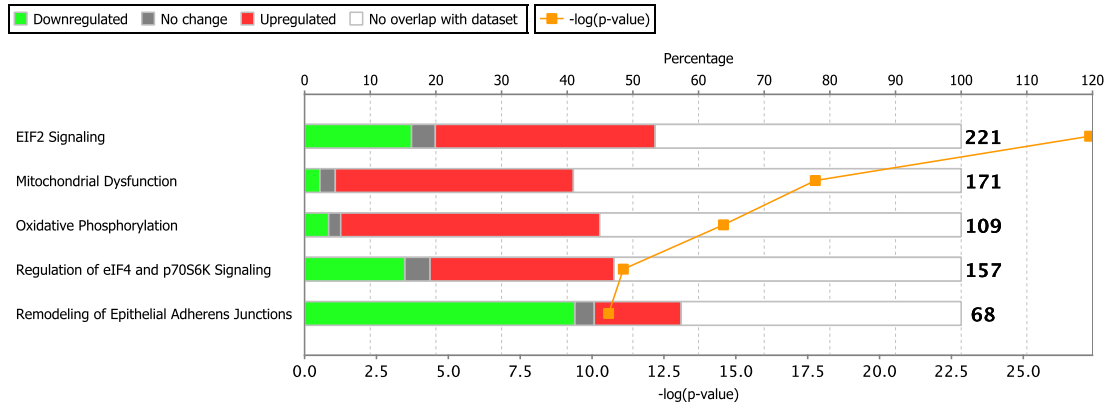




Figure 3.7: Top five canonical pathways altered in the LNCaP cell line relative to the PNT2C2 cell line. Green segments, grey segments and red segments indicate the percentage of proteins that are downregulated, unchanged, or upregulated, respectively when compared to the PNT2C2 cell line. The -log(p-value) indicates the significance of the changes of BPH-1 when compared to PNT2C2 and was calculated as a right tailed Fisher's exact test. Values at the end of the bars indicate the total number of proteins in the respective pathways.

In the PC-3 cell line 16%, 16%, 26%, 10% and 16% of proteins matched to the protein ubiquitination pathway, the EIF2 signalling pathway, the remodelling of epithelial adherens junctions pathway, the tRNA charging pathway and the epithelial adherens junction signalling pathways, respectively were downregulated when compared to PNT2C2, while 22%, 29%, 32%, 46% and 18% of the proteins were.

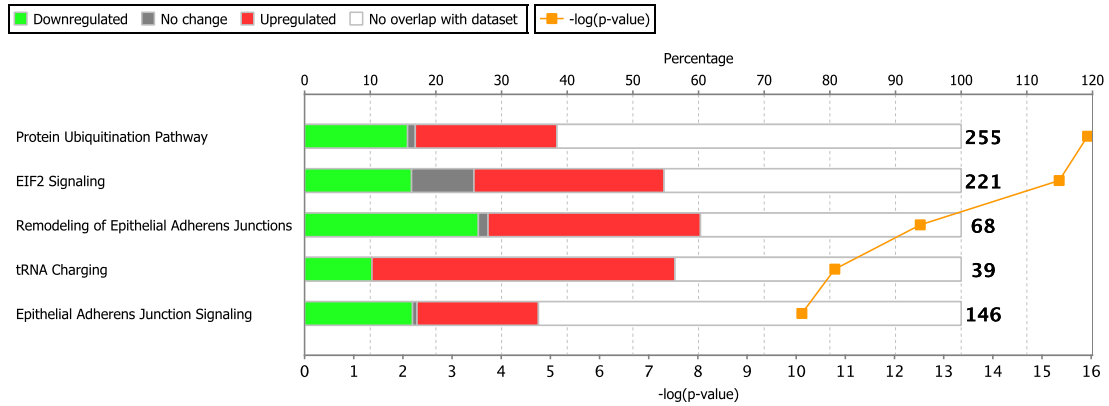




Figure 3.8: Top five canonical pathways altered in the LNCaP cell line relative to the PNT2C2 cell line. Green segments, grey segments and red segments indicate the percentage of proteins that are downregulated, unchanged, or upregulated, respectively when compared to the PNT2C2 cell line. The -log(p-value) indicates the significance of the changes of BPH-1 when compared to PNT2C2 and was calculated as a right tailed Fisher's exact test. Values at the end of the bars indicate the total number of proteins in the respective pathways.

### Upstream regulator analysis

Upstream regulators are those proteins, molecules or upstream transcriptional regulators that can explain the observed changes in the protein expression observed in the BPH-1 LNCaP and PC-3 cell lines when compared to the PNT2C2 cell line. Table 3.5, 3.6 and 3.7 list the top five upstream regulators identified by IPA for the BPH-1, LNCaP and PC-3 cell lines, respectively. The results shows that the MYC, 5-Fluorouracil, MYCN, CD-437 and TP53 molecules are the most likely upstream regulators for the BPH-1 and PC-3 cell lines, while MYC, 5-Flourouracil, MYCN, CD-437 and ST1926 are the most likely upstream regulators in the LNCaP cell line. MYC, MYCN and TP53 are proteins responsible for maintaining correct cell cycle progression, apoptosis and cellular transformation and have known interactions with multiple oncogenes such as RB1, CDKN2A/B, BRCA1, CDK12, MCM7 and CCNA2 that were in chapter 1 of this thesis.[18,75] ST1926, CD-437 and 5-Fluorouracil (FDA approved) are molecules that have shown promise as anti-cancer and anti-tumour treatments.

P-values for the BPH-1 cell line ranged from $3.27\text{x}10^{-32}$ to $1.13\text{x}10^{-50}$. Inhibition of MYC and MYCN could account for the changes seen in the BPH-1 cell line

while predicted activations were unknown for 5-Fluorouracil, CD 437 and TP53 (Table 3.5).

Table 3.5: The top five upstream regulators identified for the BPH-1 cell line.

| Identity | Name | p-value | Predicted activation |
|---|---|---|---|
| MYC | V-myc avian myelocytomatosis viral oncogene homolog | $1.13x10^{-50}$ | Inhibited |
| 5-Fluorouracil | 5-Fluorouracil | $6.42x10^{-45}$ | Unknown |
| MYCN | V-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog | $7.73x10^{-39}$ | Inhibited |
| CD-437 | CD-437 | $1.11x10^{-37}$ | Unknown |
| TP53 | Tumour protein p53 | $3.27x10^{-32}$ | Unknown |

P-values for the LNCaP cell line ranged from $1.38x10^{-35}$ to $1.10x10^{-48}$. Activation of MYC and MYCN and inhibition of 5-Fluorouracil, CD 437 and ST1926 could account for the changes seen in this cell line (Table 3.6).

Table 3.6: The top five upstream regulators identified for the LNCaP cell line.

| Identity | Name | p-value | Predicted activation |
|---|---|---|---|
| MYC | V-myc avian myelocytomatosis viral oncogene homolog | $1.10x10^{-48}$ | Activated |
| MYCN | V-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog | $9.39x10^{-40}$ | Inhibited |
| 5-Fluorouracil | 5-Fluorouracil | $3.49x10^{-38}$ | Inhibited |
| CD-437 | CD-437 | $6.85x10^{-38}$ | Activated |
| ST1926 | ST1926 | $1.38x10^{-35}$ | Inhibited |

P-values for the PC-3 cell line ranged from $2.04x10^{-28}$ to $1.58x10^{-55}$. Inhibition of 5-Fluorouracil and CD-437 could account for the changes seen in the PC-3 cell line while predicted activation were unknown for MYC, TP53 and MYCN (Table 3.7).

Table 3.7: The top five upstream regulators identified for the PC-3 cell line.

| Identity | Name | p-value | Predicted activation |
|---|---|---|---|
| MYC | V-myc avian myelocytomatosis viral oncogene homolog | $1.58x10^{-55}$ | Unknown |
| 5-Fluorouracil | 5-Fluorouracil | $1.34x10^{-40}$ | Inhibited |
| TP53 | TP53 | $9.26x10^{-38}$ | Unknown |
| CD-437 | CD-437 | $9.96x10^{-32}$ | Inhibited |
| MYCN | V-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog | $2.04x10^{-28}$ | Unknown |

### Disease analysis

The disease analysis function in IPA identifies the most likely disease related causes of the measured changes in protein expression . Table 3.8, 3.9 and 3.10 list the top five diseases which may cause the altered expression observed in the BPH-1, LNCaP and PC-3 cell lines, respectively when compared t the PNT2C cell line.. The analysis shows that cancer, organismal injury and abnormalities, tumour morphology, infectious diseases and reproductive system disease are the most likely diseases that can explain the changes in all three of the cell lines, though the ranking was slightly different between the cell lines.

P-values ranged from $5.31x10^{-05}$ to $4.28x10^{-48}$ for the BPH-1 cell line and the number of proteins identified within the disease states ranged from 125 to 1146. Interestingly cancer was identified as the top disease state in this cell line (Table 3.8).

Table 3.8: The top five diseases identified for the BPH-1 cell line.

| Name | p-value range | Number of molecules |
|---|---|---|
| Cancer | $5.31x10^{-05}$ - $4.28x10^{-48}$ | 1125 |
| Organismal injury and abnormalities | $5.31x10^{-05}$ - $4.28x10^{-48}$ | 1146 |
| Tumour morphology | $4.19x10^{-05}$ - $4.28x10^{-48}$ | 125 |
| Infectious diseases | $5.31x10^{-05}$ - $1.67x10^{-37}$ | 285 |
| Reproductive system disease | $5.31x10^{-05}$ - $1.19x10^{-23}$ | 585 |

In the LNCaP cell line the p-value ranged from $3.39x10^{-05}$ to $1.12x10^{-30}$ and the number of proteins identified within the disease states ranged from 93 to 1170. Infectious diseases ranked the highest in this cell line, followed by cancer (Table 3.9).

Table 3.9: The top five diseases identified for the LNCaP cell line.

| Name | p-value range | Number of molecules |
|---|---|---|
| Infectious diseases | $2.33x10^{-05}$ - $1.12x10^{-30}$ | 275 |
| Cancer | $3.54x10^{-05}$ - $3.06x10^{-30}$ | 1146 |
| Organismal injury and abnormalities | $3.54x10^{-05}$ - $3.06x10^{-30}$ | 1170 |
| Tumour morphology | $7.00x10^{-18}$ - $3.06x10^{-30}$ | 93 |
| Reproductive system disease | $3.39x10^{-05}$ - $1.51x10^{-19}$ | 580 |

Cancer was the top ranked disease state in the PC-3 cell line and the p-values ranged from $2.39x10^{-04}$ to $1.90x10^{-29}$ and the number of proteins identified within the disease states ranged from 91 to 1166. (Table 3.10).

Table 3.10: The top five diseases identified for the PC-3 cell line.

| Name | p-value range | Number of molecules |
|---|---|---|
| Cancer | $2.93x10^{-06}$ - $1.90x10^{-29}$ | 1148 |
| Organismal injury and abnormalities | $2.93x10^{-06}$ - $1.90x10^{-29}$ | 1166 |
| Tumour morphology | $3.27x10^{-17}$ - $1.90x10^{-29}$ | 91 |
| Infectious diseases | $5.41x10^{-07}$ - $3.43x10^{-29}$ | 269 |
| Reproductive system disease | $2.39x10^{-04}$ - $1.80x10^{-22}$ | 588 |

### Molecular and cellular function analysis

The molecular and cellular functions tool in IPA identifies those cellular functions that can most likely explain the protein expression changes in the BPH-1, LNCaP and PC-3 cell lines when compared to the PNT2C2 cell line. Table 3.11, 3.12 and 3.13 list top molecular and cellular functions for BPH-1, LNCaP and PC-3 cell lines respectively identified using IPA. The results shows that cell death and survival, cellular growth and proliferation, RNA post-transcriptional modification, protein synthesis and cellular development are the most likely molecular and cellular events that can explain the changes in the BPH-1, LNCaP and PC-3 cell lines when compared to PNT2C2. Cell death and survival was the top ranked molecular and cellular function identified in all three cell lines.

The p-values for the BPH-1 cell line ranged from $3.61x10^{-05}$ to $4.28x10^{-48}$ and the number of proteins identified within the molecular and cellular functions ranged from 102 to 544. (Table 3.11).

Table 3.11: The top five molecular and cellular functions identified for the BPH-1 cell line.

| Name | p-value range | Number of molecules |
|---|---|---|
| Cell death and survival | $4.96x10^{-05}$ - $4.28x10^{-48}$ | 544 |
| Cellular growth and proliferation | $3.62x10^{-06}$ - $3.08x10^{-41}$ | 536 |
| RNA post-transcriptional modification | $5.31x10^{-05}$ - $2.12x10^{-40}$ | 102 |
| Protein synthesis | $3.36x10^{-06}$ - $1.91x10^{-26}$ | 222 |
| Cellular development | $3.61x10^{-05}$ - $3.48x10^{-22}$ | 278 |

In the LNCaP cell line the p-values ranged from $1.55x10^{-05}$ to $1.11x10^{-37}$ and the number of proteins identified within the molecular and cellular functions ranged from 94 to 535. (Table 3.12).

Table 3.12: The top five molecular and cellular functions identified for the LNCaP cell line.

| Name | p-value range | Number of molecules |
|---|---|---|
| Cell death and survival | $3.45x10^{-05}$ - $1.11x10^{-37}$ | 535 |
| RNA post-transcriptional modification | $2.34x10^{-06}$ - $4.44x10^{-33}$ | 94 |
| Cellular growth and proliferation | $1.55x10^{-05}$ - $2.665x10^{-30}$ | 514 |
| Protein synthesis | $2.26x10^{-05}$ - $2.39x10^{-25}$ | 232 |
| Cellular development | $1.55x10^{-05}$ - $1.79x10^{-19}$ | 281 |

The p-values ranged from $2.28x10^{-06}$ to $3.49x10^{-52}$ in the PC-3 cell line and the number of proteins identified within the molecular and cellular functions ranged from 94 to 568. (Table 3.13)

Table 3.13: The top five molecular and cellular functions identified for the PC-3 cell line.

| Name | p-value range | Number of molecules |
|---|---|---|
| Cell death and survival | $9.36x10^{-07}$ - $3.49x10^{-52}$ | 568 |
| Cellular growth and proliferation | $2.28x10^{-06}$ - $3.38x10^{-48}$ | 555 |
| RNA post-transcriptional modification | $5.03x10^{-09}$ - $5.51x10^{-35}$ | 94 |
| Protein synthesis | $5.52x10^{-07}$ - $1.36x10^{-29}$ | 241 |
| Cellular development | $2.28x10^{-06}$ - $4.27x10^{-21}$ | 273 |

### GO term analysis

The analysis of GO term annotations for biological process, cellular location and molecular function can provide further insights into the distribution and activity of the proteins identified for the individual cell lines. This can aid the understanding and interpretation of large data sets and provide information on local and integrated functions throughout the cell lines.[162,163] The use of pie and bar charts can allow qualitative and semi-quantitative interpretation the biological significance of proteomics data.[158] Biological processes, cellular location and molecular function information based on GO term annotations were retrieved from the NCBI database for all proteins identified and the proportions of each GO term was calculated in Scaffold and is listed in figure 3.9, 10 and 3.11, respectively below. It should be noted that proteins may possess more than one GO term annotation and can therefore be counted in multiple categories.

The relative proportions of proteins assigned to biological process GO terms are shown in figure 3.9. Large proportions of identified proteins were assigned to cellular process (18.48%) metabolic process (15.71%) and biological regulation (10.86%). A smaller proportion (1-10%) of proteins were assigned to localisation, response to stimulus, development process, multicellular organismal process, viral reproduction, reproduction, reproductive process and multi-organism process. The percentage of proteins assigned to locomotion, biological adhesion, rhythmic process, growth, cell killing were under 1%. Of the biological processes 14.59% were unknown.

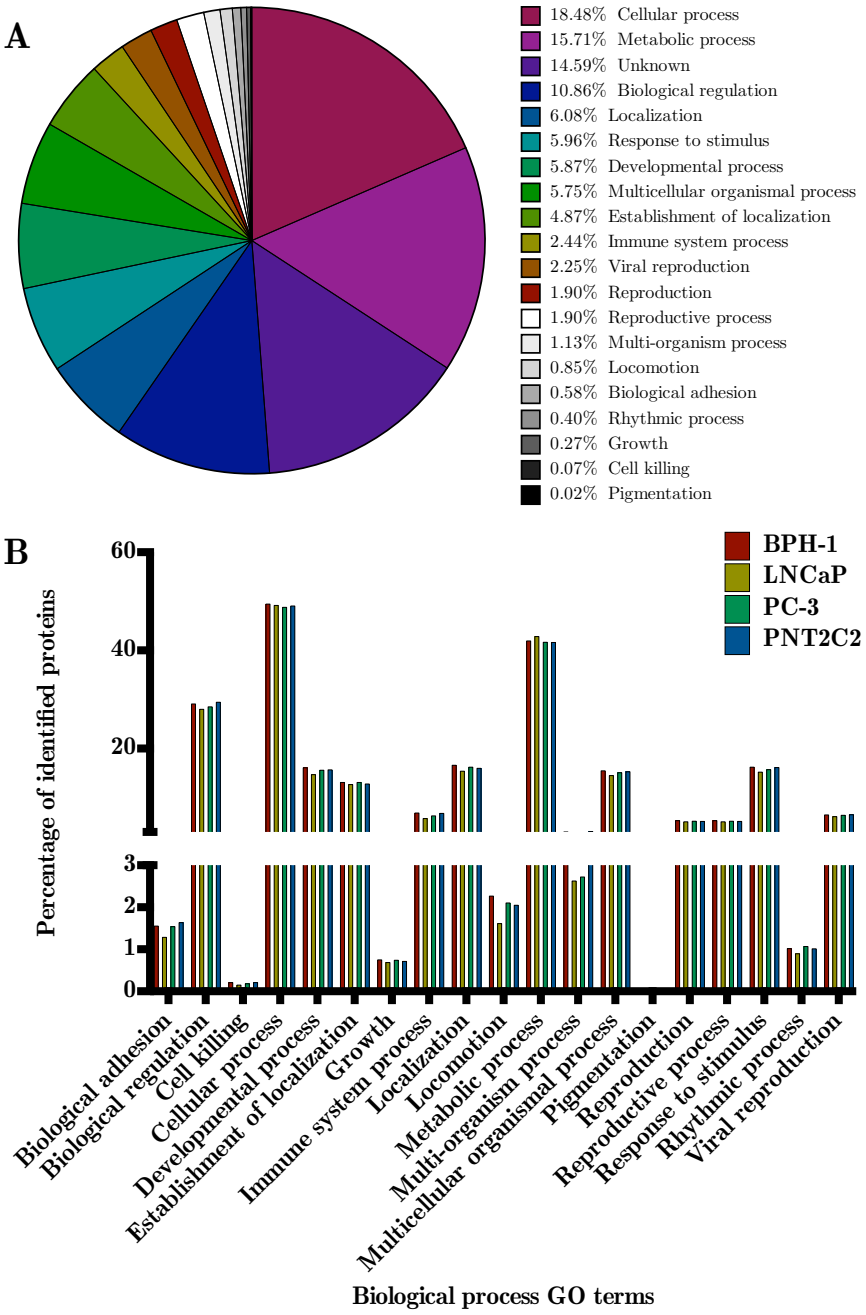


Figure 3.9: The relative proportions of proteins that were assigned biological process GO terms for the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines (A) and separately for the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines respectively (B). GO term annotations were retrieved from the NCBI database. Proteins may possess multiple GO annotations.

The relative proportions of proteins assigned to cellular component GO terms are shown in figure 3.10. Large proportions of identified proteins were assigned to intracellular organelle (15.95%), cytoplasm (14.32%), organelle part (12.55%) and the nucleus (10.49%). A smaller proportion (10-1%) of proteins were assigned to the membrane, extracellular region, plasma membrane, mitochondrion, organelle membrane, cytoskeleton, endoplasmic reticulum, ribosome and Golgi apparatus. The percentage of proteins assigned to the endosome was under 1%. 12.34% of the proteins could not be assigned.

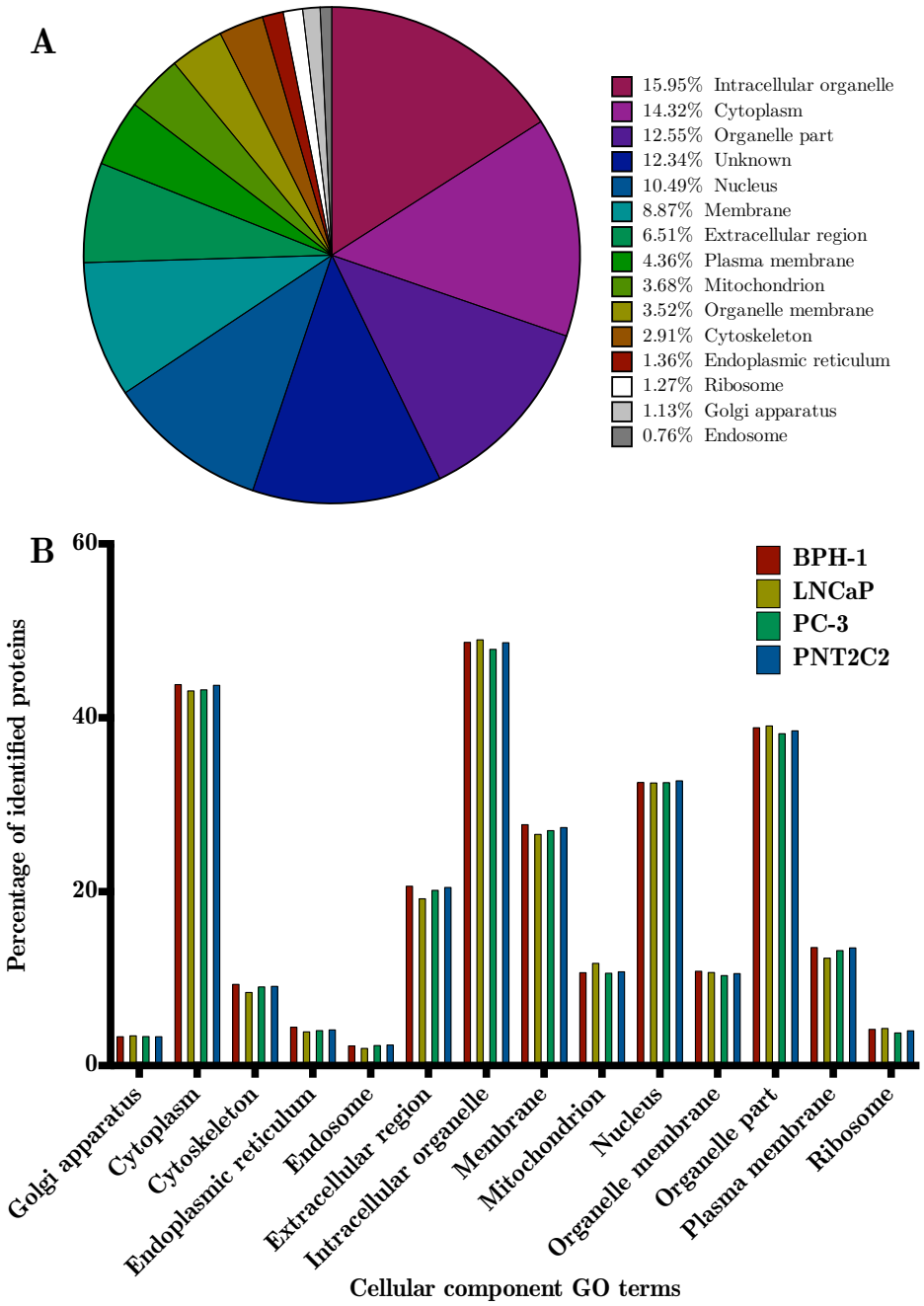


Figure 3.10: The relative proportions of proteins that were assigned cellular components GO terms for the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines (A) and separately for the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines respectively (B). GO term annotations were retrieved from the NCBI database. Proteins may possess multiple GO annotations.

The relative proportions of proteins assigned to molecular function GO terms are shown in figure 3.11. Large proportions of identified proteins were assigned to molecular function (30.91%), binding (27.31%), unknown (18.99%) and catalytic activity (12.52%). A smaller proportion (10 - 1%) of proteins were assigned to structural molecule activity, enzyme regulator activity, transcription regulator activity and transporter activity. The percentage of proteins assigned to molecular transducer activity, translation regulator activity, electron carrier activity, motor activity, auxiliary transport protein activity, chemoattractant activity, metallochaperone activity and protein tag were under 1%. A total of 18.99% of proteins could not be assigned.

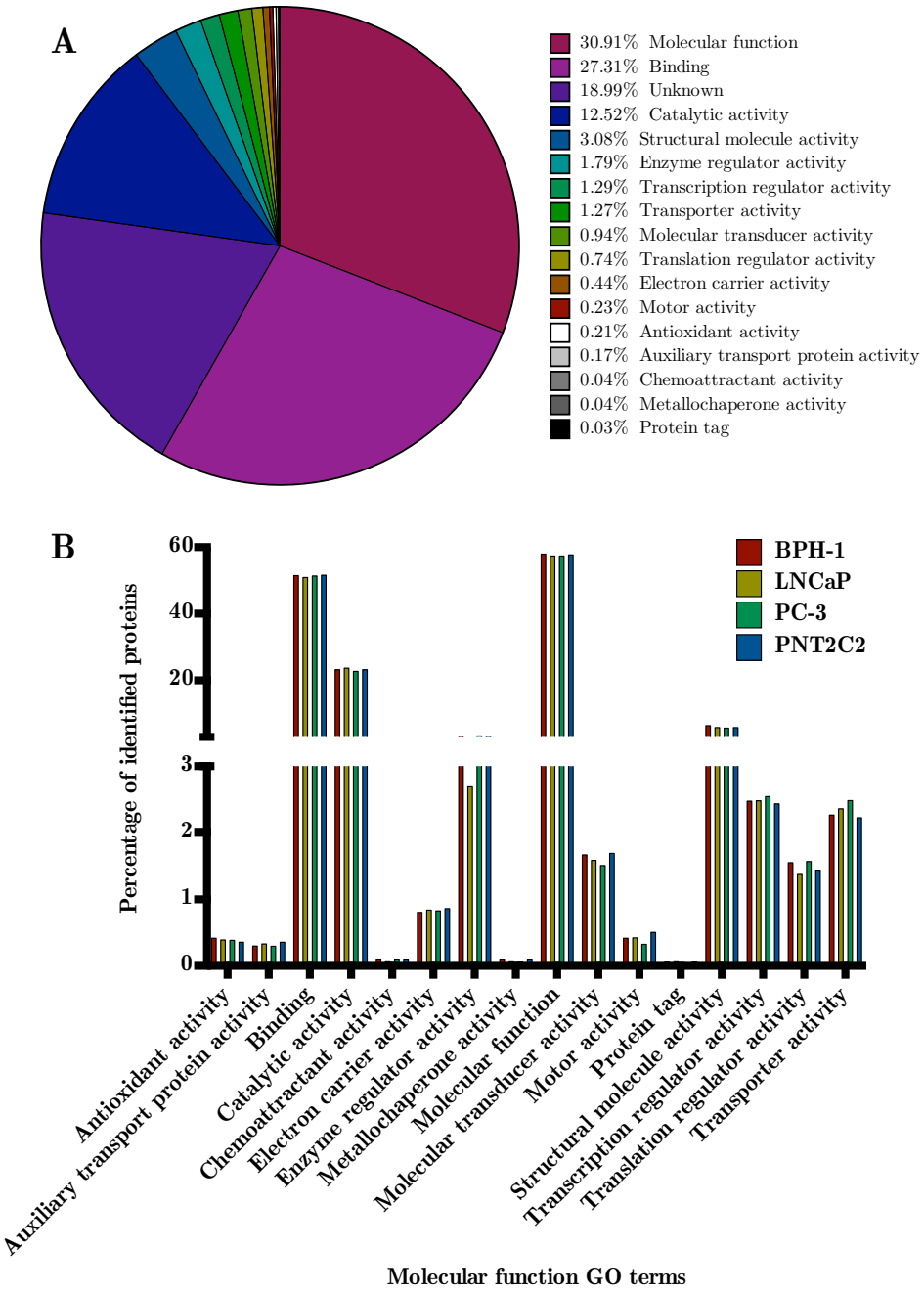


Figure 3.11: The relative proportions of proteins that were molecular function GO terms for the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines (A) and separately for the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines respectively (B). GO term annotations were retrieved from the NCBI database. Proteins may possess multiple GO annotations.

### Candidate biomarker selection

Membrane bound and extracellular proteins have previously been used as the best possible candidate biomarkers as they are easily secreted into serum and can be analysed non invasively.[164] Scaffold was utilised to retrieve GO terms from the NCBI database and identify membrane and extracellular proteins from the BPH-1, LNCaP, PC-3 and PNT2C2 whole cell proteome. A total of 1176 proteins containing membrane bound and/or extracellular GO term annotations were identified. Of these 544 (46%) were common to all cell lines, 56 (5%), 93 (8%), 59 (5%) and 58 (5%) proteins were unique to BPH-1, LNCaP, PC-3 and PNT2C2, respectively. The overlap of identified proteins is shown below in the Venn diagram (figure 3.12).

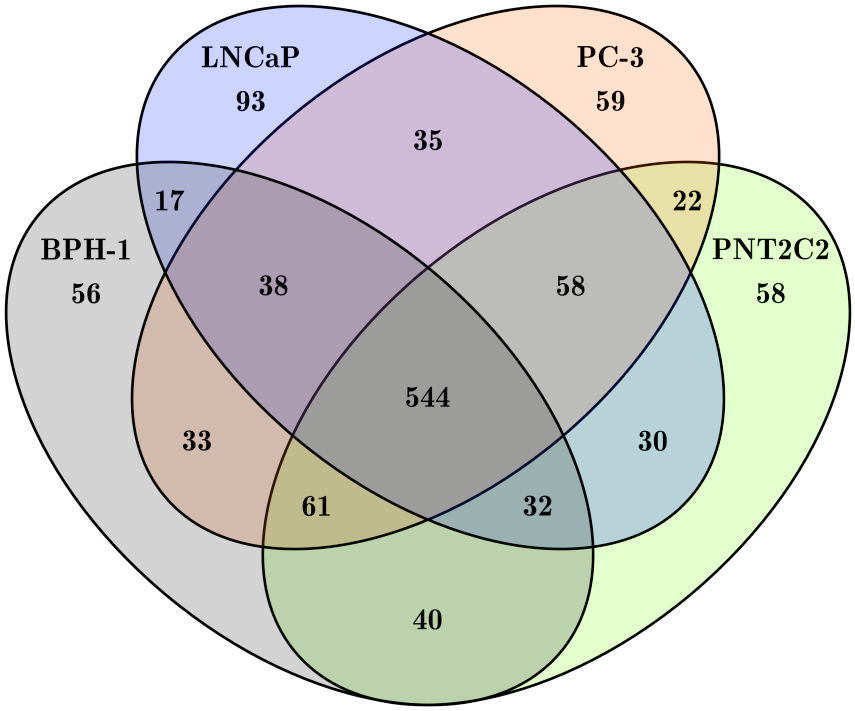


Figure 3.12: Venn diagram showing the overlap of membrane bound and/or extracellular proteins identified in the proteome of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines.

The analysis of this data is useful for narrowing down proteins which may be useful as candidate biomarkers unique to prostate cancer. The membrane and extracellular proteins which were unique to the prostate cancer cell lines, LNCaP and PC-3, or which occurred in both LNCaP and PC-3 cells, but not BPH-1 or PNT2C2 cells were therefore identified and are listed in supplementary tables 5.1 -

5.3. These proteins represent potential candidates biomarkers for PCa, which can be investigated further.

Several candidate biomarkers were identified that were unique to the LNCaP and PC-3 cell lines. The biomarker filter function was used to identify interesting proteins and extracellular and plasma membrane proteins that were unique to the LNCaP and PC-3 cell lines were selected. Table 3.14 and 3.15 lists a number of proteins which may act as candidate biomarkers identified in the whole cell samples.

Table 3.14: Candidate biomarkers from the proteome for the LNCaP cell line identified using IPA.

| Entry | Entry name | Protein names | Gene names | Expr p-value |
|---|---|---|---|---|
| Q12955 | ANK3_HUMAN | Ankyrin-3 (ANK-3) (Ankyrin-G) | *ANK3* | $9.20\text{x}10^{-03}$ |
| C9JJX6 | C9JJX6_HUMAN | Armadillo repeat protein deleted in velo-cardio-facial syndrome | *ARVCF* | $3.70\text{x}10^{-02}$ |
| Q04609 | FOLH1_HUMAN | Glutamate carboxypeptidase 2 (EC 3.4.17.21) (Cell growth-inhibiting gene 27 protein) (Folate hydrolase 1) (Folylpoly-gamma-glutamate carboxypeptidase) (FGCP) (Glutamate carboxypeptidase II) (GCPII) (Membrane glutamate carboxypeptidase) (mGCP) (N-acetylated-alpha-linked acidic dipeptidase I) (NAALADase I) (Prostate-specific membrane antigen) (PSM) (PSMA) (Pteroylpoly-gamma-glutamate carboxypeptidase) | *FOLH1 FOLH NAALAD1 PSM PSMA GIG27* | $6.80\text{x}10^{-04}$ |
| P07288 | KLK3_HUMAN | Prostate-specific antigen (PSA) (EC 3.4.21.77) (Gamma-seminoprotein) (Seminin) (Kallikrein-3) (P-30 antigen) (Semenogelase) | *KLK3 APS* | $1.00\text{x}10^{-04}$ |
| Q9BRT6 | LLPH_HUMAN | Protein LLP homolog (Protein LAPS18-like) | *LLPH C12orf31 cPERP-G* | $4.90\text{x}10^{-03}$ |
| Q9NX58 | LYAR_HUMAN | Cell growth-regulating nucleolar protein | *LYAR PNAS-5* | $1.00\text{x}10^{-04}$ |
| Q15599 | NHRF2_HUMAN | Na(+)/H(+) exchange regulatory cofactor NHE-RF2 (NHERF-2) (NHE3 kinase A regulatory protein E3KARP) (SRY-interacting protein 1) (SIP-1) (Sodium-hydrogen exchanger regulatory factor 2) (Solute carrier family 9 isoform A3 regulatory factor 2) (Tyrosine kinase activator protein 1) (TKA-1) | *SLC9A3R2 NHERF2* | $1.00\text{x}10^{-04}$ |

Table 3.15: Candidate biomarkers from the proteome for the PC-3 cell line identified using IPA.

| Entry | Entry name | Protein names | Gene names | Expr p-value |
|---|---|---|---|---|
| P48960 | CD97_HUMAN | CD97 antigen (Leukocyte antigen CD97) (CD antigen CD97) [Cleaved into: CD97 antigen subunit alpha; CD97 antigen subunit beta] | *CD97* | $3.40x10^{-02}$ |
| Q9Y2D5 | AKAP2_HUMAN | A-kinase anchor protein 2 (AKAP-2) (AKAP-KL) (Protein kinase A-anchoring protein 2) (PRKA2) | *AKAP2 KIAA0920 PRKA2* | $6.60x10^{-04}$ |
| P08133 | ANXA6_HUMAN | Annexin A6 (67 kDa calelectrin) (Annexin VI) (Annexin-6) (Calphobindin-II) (CPB-II) (Chromobindin-20) (Lipocortin VI) (Protein III) (p68) (p70) | *ANXA6 ANX6* | $1.00x10^{-04}$ |
| F8VUA2 | F8VUA2_HUMAN | Charged multivesicular body protein 1a | *CHMP1A* | $1.90x10^{-03}$ |
| Q9Y696 | CLIC4_HUMAN | Chloride intracellular channel protein 4 (Intracellular chloride ion channel protein p64H1) | *CLIC4* | $1.00x10^{-04}$ |
| P09496 | CLCA_HUMAN | Clathrin light chain A (Lca) | *CLTA* | $1.00x10^{-04}$ |
| Q9NQX3 | GEPH_HUMAN | Gephyrin [Includes: Molybdopterin adenylyltransferase (MPT adenylyltransferase) (EC 2.7.7.75) (Domain G); Molybdopterin molybdenumtransferase (MPT Mo-transferase) (EC 2.10.1.1) (Domain E)] | *GPHN GPH KIAA1385* | $1.00x10^{-04}$ |
| P23229 | ITA6_HUMAN | Integrin alpha-6 (CD49 antigen-like family member F) (VLA-6) (CD antigen CD49f) [Cleaved into: Integrin alpha-6 heavy chain; Integrin alpha-6 light chain; Processed integrin alpha-6 (Alpha6p)] | *ITGA6* | $1.00x10^{-04}$ |
| Q9BRT6 | LLPH_HUMAN | Protein LLP homolog (Protein LAPS18-like) | *LLPH C12orf31 cPERP-G* | $1.00x10^{-04}$ |
| Q9NX58 | LYAR_HUMAN | Cell growth-regulating nucleolar protein | *LYAR PNAS-5* | $1.00x10^{-04}$ |
| P48681 | NEST_HUMAN | Nestin | *NES Nbla00170* | $1.00x10^{-04}$ |
| Q8TEW0 | PARD3_HUMAN | Partitioning defective 3 homolog (PAR-3) (PARD-3) (Atypical PKC isotype-specific-interacting protein) (ASIP) (CTCL tumor antigen se2-5) (PAR3-alpha) | *PARD3 PAR3 PAR3A* | $1.00x10^{-04}$ |
| O00264 | PGRC1_HUMAN | Membrane-associated progesterone receptor component 1 (mPR) | *PGRMC1 HPR6.6* | $4.10x10^{-02}$ |
| Q13835 | PKP1_HUMAN | Plakophilin-1 (Band 6 protein) (B6P) | *PGRMC PKP1* | $1.00x10^{-04}$ |
| Q5F2F8 | Q5F2F8_HUMAN | Serine/threonine-protein phosphatase (EC 3.1.3.16) | *PPP3CB hCG_18332* | $1.30x10^{-02}$ |
| Q14558 | KPRA_HUMAN | Phosphoribosyl pyrophosphate synthase-associated protein 1 (PRPP synthase-associated protein 1) (39 kDa phosphoribosypyrophosphate synthase-associated protein) (PAP39) | *PRPSAP1* | $3.40x10^{-02}$ |
| Q5QPE2 | Q5QPE2_HUMAN | Protein transport protein Sec23B (Fragment) | *SEC23B* | $3.80x10^{-03}$ |
| H3BN50 | H3BN50_HUMAN | Na(+)/H(+) exchange regulatory cofactor NHE-RF2 (Fragment) | *SLC9A3R2* | $1.70x10^{-02}$ |
| Q9Y6W5 | WASF2_HUMAN | Wiskott-Aldrich syndrome protein family member 2 (WASP family protein member 2) (Protein WAVE-2) (Verprolin homology domain-containing protein 2) | *WASF2 WAVE2* | $3.60x10^{-02}$ |

## 3.4 Secretome analysis

Following the optimisation of seeding densities for the production of secretome samples (Section 3.1) and the optimisation of the protein precipitation method (Section 3.2) three independent biological replicates were collected and prepared for analysis by MS. Unfortunately, the due to technical difficulties experienced resulting from construction in the building in which the MS is housed the MS had to be taken off line and will only be brought online following the completion of the renovations in 2017. As a result these samples could not be analysed in time for the completion of this MSc thesis.

Preliminary data secretome data used to test the proof of concept was, however, acquired at an earlier date. It should, however, be noted that while this data was acquired using the same criteria as the samples described above, they only included a single biological replicate, and the proteins were precipitated using ammonium sulphate prior to analysis.

Scaffold was utilised to obtain protein identities for the BPH-1, LNCaP, PC-3 and PNT2C2 secretome samples as previously described for the proteome (section 3.3). A total of 1106 proteins were identified. Of these 310 (28%) were common to all cell lines, 34 (3%), 266 (24%), 103 (9%) and 46 (4%) proteins were independent to BPH-1, LNCaP, PC-3 and PNT2C2, respectively as shown in the Venn diagram below (figure 3.13).

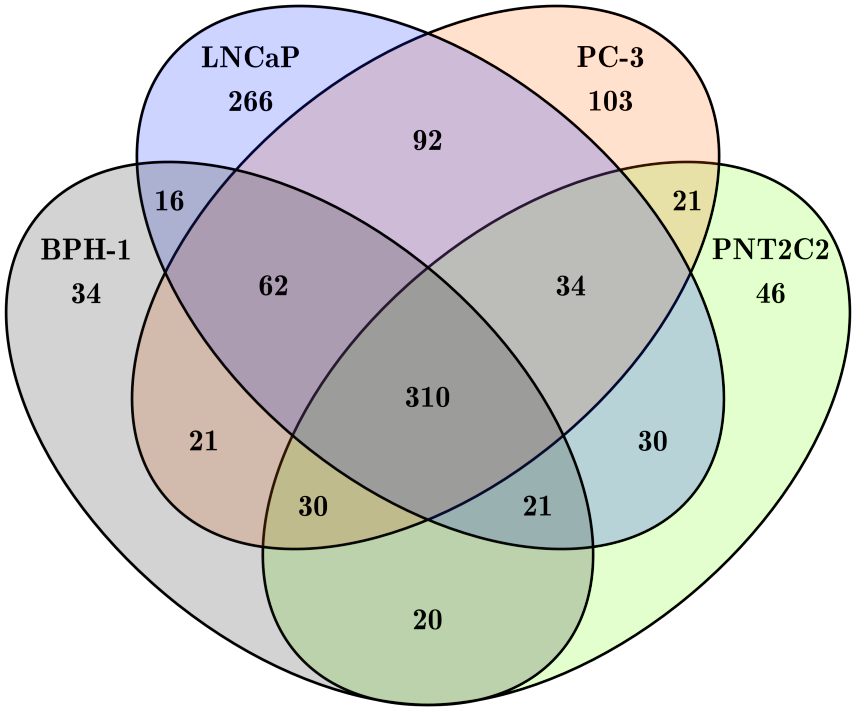


Figure 3.13: Venn diagram showing the overlap of all proteins identified in the secretome of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines.

As previously stated membrane bound and extracellular proteins have previously been used as the best possible candidate biomarkers as they are easily secreted into serum and can be analysed non invasively.[164] Scaffold was utilised to retrieve GO terms from the NCBI database and identify membrane and extracellular proteins from the BPH-1, LNCaP, PC-3 and PNT2C2 secretome. A total of 393 proteins containing membrane bound and/or extracellular GO term annotations were identified using the NCBI database, of these 147 (37%) were common to all cell lines, 5 (1%), 67 (17%), 31 (8%) and 10 (3%) proteins were unique to BPH-1, LNCaP, PC-3 and PNT2C2, respectively. The overlap of identified proteins is shown below in the Venn diagram (figure 3.14).

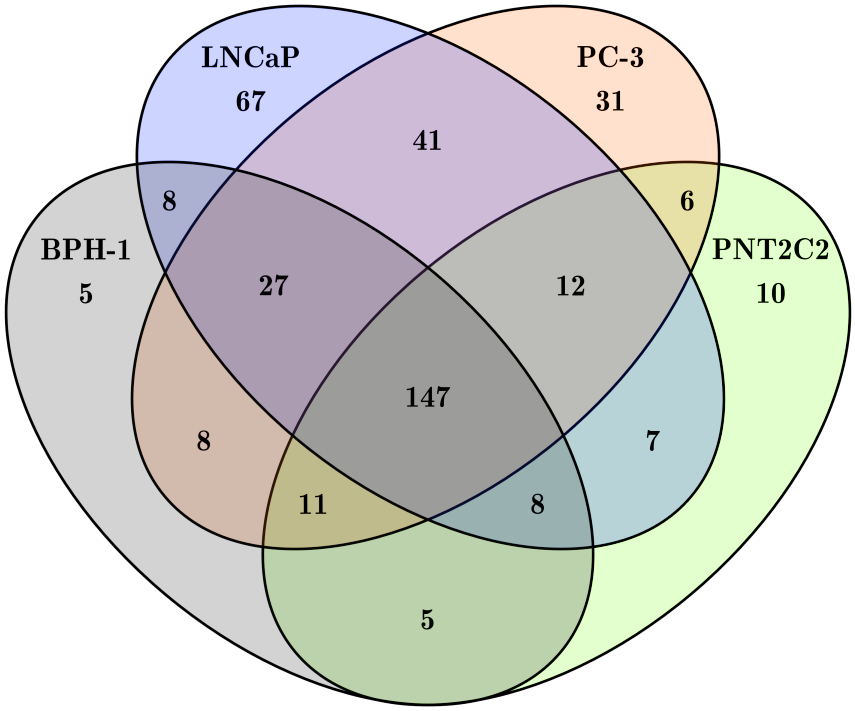


Figure 3.14: Venn diagram showing the overlap of membrane bound and/or extracellular proteins identified in the secretome of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines.

The analysis of this data is useful for narrowing down proteins which may be useful as candidate biomarkers unique to prostate cancer. The membrane and extracellular proteins with were unique to the prostate cancer cell lines, LNCaP and PC-3, or which occurred in both LNCaP and PC-3 cells, but not BPH-1 or PNT2C2 cells were therefore identified and are listed in supplementary tables 5.4 - 5.6. Proteins that have previously been reported as potential biomarkers were identified are listed in table 3.16 and 3.17. These proteins therefore represent candidates which can be investigated further as biomarkers for PCa.

Table 3.16: Candidate biomarkers for prostate cancer identified from the secretome using IPA.

| Entry | Entry name | Protein names | Gene names | Expr p-value |
|---|---|---|---|---|
| Q12955 | ANK3_HUMAN | Ankyrin-3 (ANK-3) (Ankyrin-G) | *ANK3* | $7.20 \times 10^{-03}$ |
| P54105 | ICLN_HUMAN | Methylosome subunit pICln (Chloride channel, nucleotide sensitive 1A) (Chloride conductance regulatory protein ICln) (I(Cln)) (Chloride ion current inducer protein) (ClCI) (Reticulocyte pICln) | *CLNS1A CLCI ICLN* | $3.40 \times 10^{-02}$ |
| Q00610 | CLH1_HUMAN | Clathrin heavy chain 1 (Clathrin heavy chain on chromosome 17) (CLH-17) | *CLTC CLH17 CLTCL2 KIAA0034* | $3.90 \times 10^{-03}$ |
| P05161 | ISG15_HUMAN | Ubiquitin-like protein ISG15 (Interferon-induced 15 kDa protein) (Interferon-induced 17 kDa protein) (IP17) (Ubiquitin cross-reactive protein) (hUCRP) | *ISG15 G1P2 UCRP* | $9.60 \times 10^{-03}$ |
| P26038 | MOES_HUMAN | Moesin (Membrane-organizing extension spike protein) | *MSN* | $9.90 \times 10^{-03}$ |
| B4DH47 | B4DH47_HUMAN | cDNA FLJ53769, highly similar to Homo sapiens nebulette (NEBL), transcript variant 2, mRNA | N/A | $2.30 \times 10^{-02}$ |
| P48681 | NEST_HUMAN | Nestin | *NES Nbla00170* | $3.80 \times 10^{-03}$ |
| Q13813 | SPTN1_HUMAN | Spectrin alpha chain, non-erythrocytic 1 (Alpha-II spectrin) (Fodrin alpha chain) (Spectrin, non-erythroid alpha subunit) | *SPTAN1 NEAS SPTA2* | $4.20 \times 10^{-03}$ |
| Q01082 | SPTB2_HUMAN | Spectrin beta chain, non-erythrocytic 1 (Beta-II spectrin) (Fodrin beta chain) (Spectrin, non-erythroid beta chain 1) | *SPTBN1 SPTB2* | $1.20 \times 10^{-04}$ |

# Chapter 4

# Discussion

## 4.1 Introduction

PCa is the second most common cause of cancer internationally and has an estimated mortality rate of 6.6%.[23] PSA has been the industry standard biomarker and test for PCa since its discovery in 1971 and approval by the FDA in 1986.[113,165,166] PSA was initially thought to be an effective test for PCa that could be obtained non-invasively through a blood test.[106] In 1994 the FDA approved the simultaneous use of the DRE in asymptomatic men.[167] Subsequently, PSA was found to be increased in patients with more benign conditions such as BPH and prostatitis.[114] The large incidences of false positives, misdiagnosis and the lack of tangible benefits in terms of reducing mortality of the PSA/DRE further decreased the confidence in its diagnostic potential.[105,106,114] PSA has since been widely criticised and discredited with the initial discoverer of the PSA test Dr Richard Ablin stating that it is not a reliable indicator of PCa.[168] Subsequently the U.S. Preventive Services Task Force in agreement with the FDA recommended that PSA not be used as screening methods for PCa.[169]

As a result constant research has been undertaken to develop novel non-invasive biomarkers that could not only identify PCa but also indicate the severity and possible progression of the disease and give an indication of the most beneficial treatments.[75] The introduction of the PHI which utilises PSA, proPSA and fPSA and mathematical modelling has significantly increased the efficacy of PCa diagnosis.[104] Several new promising biomarkers are currently in development including PCA3, TMPRSS2:ERG, and the combinational use of several biomarkers at once such as ProMARK and ConfirmMDx.[170]

Prostate cancer is a complex and variable disease and the utilisation of clinical samples does not usually provide unbiased reliable biomarkers without an extremely large sample size.[171] This study contributed to the search for novel biomarkers

by investigating differences in cancerous, benign and normal prostate tissue cell lines. It was hypothesised that there may be distinct protein profiles in terms of protein expression levels and unique proteins that could act as potential diagnostic tools for PCa. The primary goal of this study was therefore to identify potential biomarker candidates and to confirm the presence of candidate biomarkers identified by previous studies by characterising the proteome and secretome of the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines. The culturing conditions of the four lines were therefore optimised for mass spectrometric analysis and an in depth analysis of the proteome and secretome profiles were performed.

## 4.2 Cell culture optimisation

Accurate determination of biomarkers in secretome samples dictates that the presences of intracellular contaminating proteins released during cell death is minimised.[172] This was achieved by optimising the initial seeding densities (11400 - 126 000 cells per $cm^2$) of the cell lines to increase protein secretion while minimising cell death. Furthermore, a chemically defined serum free media (CDCHO) was used to reduce the presence of interfering proteins.

By measuring the ratio of LDH to protein concentration the lowest seeding densities for BPH-1 and PC-3 cell lines and the highest seeding densities for LNCaP and PNT2C2 cell lines were shown to be the most optimal for mass spectrometric analysis as seen in figure 3.3. Previous work by Makridakis et al., Sardana et al., Kulasingam et al. and Gunawardana et al. has shown that lower the seeding densities tend to be more optimal for mass spectrometric analysis.[153,173–175] However, while Sardana et al. observed this trend in PC-3 cells, the same trend was not observed in LNCaP cells therefore suggesting that the optimal seeding density is not necessarily the lowest in all cases but may be linked to the cell line in question.[153]

While the higher seeding densities tended to result in the lowest LDH/protein ratio for the LNCaP and PNT2C2 cell lines in this study, it should be noted that the determined ratios were not significantly different from those obtained at other lower seeding densities (Figure 3.3). It was therefore also important to consider the absolute LDH concentration, as this is an indication of cell lysis, which could potentially complicate the analysis of the secretome. The lowest seeding densities were therefore used for secretome analysis in all four cell lines as these tended to contain the lowest LDH levels and adequate protein concentrations. Although it was tempting to reduce the seeding density even lower than the range tested in this study Sardana et al. and Peh et al. have previously shown that this had no significant benefit on cell culture optimisation.[153,176] Conversely, Peh et al. showed that unfavourable cell morphology and dispersion was significantly higher at very

low seeding densities (2500 and 5000 cells per $cm^2$). Furthermore, Sardana et al. found that medium seeding densities in the LNCaP (85 700 cells per $cm^2$) resulted in higher LDH concentrations compared to higher seeding densities (126 000 cells per $cm^2$). Sardana et al also showed that low seeding densities (10 900 - 42 900 cells per $cm^2$) in the PC-3 cell line reach a limit whereby decreasing the seeding density does not result in decreased cell death.[153]

While cell culture conditions can be optimised to minimise cell lysis, this cannot be eliminated completely. This study therefore employed an additional approach, whereby the proteome of each cell line was also characterised. This proteome profile containing possible contaminant intracellular proteins was subsequently used as a reference when investigating the secretome in order to improve the likelihood of identifying candidate biomarkers which were true secreted proteins. It should, however, be noted that while this approach can aid in interpreting the ambiguity brought about intracellular leakage it does not eliminate the effect of the intracellular proteins on the dynamic range measured by the MS and thus minimising these proteins therefore remains important.

In summary, the results obtained in this study confirm the trend observed previously that low but not extremely low seeding densities tend to produce a more optimal LDH to protein concentration ratio. The approach used in this study confirmed the optimal seeding densities for the LNCaP and PC-3 cell lines previously identified by Sardana et al. and led to the successful identification of optimal seeding densities for the BPH-1 and PNT2C2 cell lines.

## 4.3 Protein precipitation

Mass spectrometry is inherently a concentration dependent technique.[121] A bottom up approach is very sensitive to individual peptide concentration and this has an effect on protein identification and subsequent quantification.[121,123] Therefore, it is important to increase the concentration of proteins prior to mass spectrometry to significantly increase not only the number of the proteins identified but also the significance of the identities.

Even though mass spectrometry based proteomics is a very sensitive technique, the identification and quantification of proteins from secretome samples requires a step to concentrate the proteins prior to analysis.[177] Secretome samples obtained from the collection of CDCHO media from tissue culture samples were therefore subjected to ammonium sulphate, acetone, methanol/chloroform and PTA-mediated acetone precipitation in order to determine which precipitation method would yield the best recovery.

Ammonium sulphate precipitation, also known as salting out, reduces the hydration layer, reducing surface charges increasing hydrophobic interactions between

proteins and initiates protein-protein interactions which forces mostly hydrophobic proteins out of solution.[178] Unfortunately, the additional salt which needs to be removed, usually using dialysis, can result in unwanted protein losses.[179] Acetone precipitation exploits the difference of dielectric constants in water and proteins.[179] The high dielectric constant of water allows it to have stable interactions with proteins.[179] These interactions can be reduced through the addition of an organic solvent with a low dielectric constant such as acetone.[179] Acetone reduces the dispersion of proteins and lowers the effective dielectric constant of water allowing aggregation of proteins.[179] As this method reduces the hydrophobic effect it results in proteins denaturing, which can make them difficult to re-solubilise.[179,180] Nonetheless, acetone is an effective means of sample preparation allowing removal of contaminants while simultaneously concentrating proteins.[179] Methanol chloroform utilises a phase separation principle to trap and concentrate proteins between an organic and inorganic layer of solvents.[179] The addition of methanol causes proteins to form a hydrophobic and hydrophilic layer at the interphase.[179] Subsequent removal of the aqueous layer and the addition of chloroform allows the collection of the proteins.[179] It acts an effective means of collecting proteins free of contaminants.[179] Unfortunately, not all proteins are precipitated as this method favours hydrophobic proteins.[179] PTA acetone precipitation works on a combinational approach using the acetone to precipitate hydrophobic proteins and the PTA facilitating further precipitation of polar proteins through negative ion precipitation.[181] The addition of PTA to an acidic media solution results in positively charged proteins binding to PTA resulting in the precipitation of the proteins in their salt form.[182] This in an effective means of recovering a high concentration of diverse proteins.[181]

The results obtained in this study (Table 3.1 and Figure 3.4) show that acetone precipitation resulted in the highest recovery (56%). The recovery obtained using ammonium sulphate precipitation was substantially lower (35%) and no protein was recovered when using the methanol chloroform method (0%). It should be noted that an interphase of possible proteins was observed during methanol chloroform precipitation but, could not be recovered. Furthermore, as this method favours the purification of hydrophobic proteins, it is likely that more polar water soluble secreted proteins would be easily lost during methanol chloroform precipitation.[183] The protein concentration from the PTA-mediated acetone precipitation could not be accurately determined and this method was thus excluded. This is likely due to the use of the Bradford assay which was chosen as it is more robust in determining protein concentrations in the presence of interfering agents,[184] but which can affected by the relative pH of the samples.[151] The addition of PTA may therefore have accounted for inconsistent results when using the Bradford assay. The presence of large precipitates that were difficult to re-solubilise post-precipitation may also have interfered with spectrophotometric measurements,

thereby compromising the reliability of the assay.

The observation that the acetone precipitation method yielded the largest recoveries is in agreement with previous investigations by Yuan et al., Duan et al. and Sickmann et al. which showed that acetone precipitation was a superior method when compared to other sample preparation methods such as ultrafiltration, dialysis, TCA/acetone precipitation, ethanol precipitation, bio-spin columns and strong cation exchange chromatography.[185–187] Moreover, the use of acetone precipitation allows for the increased recovery of smaller proteins, which get trapped in the hydrophobic space between larger proteins.[179] Acetone precipitation therefore allows for superior recovery due to decreased handling and expedited processing as no sample clean-up is necessary. While the recovery achieved in this study was lower than that reported by Yuan et al. (94%) it was substantially higher than that achieved by Sickmann et al. (40-50%) and Sardana et al. (17-25%).

## 4.4 Proteome analysis

The recent advancements in mass spectrometry have allowed rapid progression of cost effective, high throughput meaningful proteomic studies.[188] This study utilised four prostate epithelial cell lines, namely BPH-1, LNCaP, PC-3 and PNT2C2. BPH-1 is benign hyperplastic prostatic epithelial cell line obtained during a transurethral resection and immortalised with SV40 that gives a good indication of benign uncontrolled growth of prostate tissue.[148] The LNCaP cell line obtained from a human metastatic prostate adenocarcinoma localised to the lymph node has been a staple of multiple studies of early PCa due to its dependence on androgen stimulation.[146] It has since been found that LNCaP cells are not intrinsically tumorigenic without the introduction of several external components such as the extracellular matrix and paracrine-mediated growth factors highlighting the requirement for several factors to maintain metastatic ability.[189,190] PC-3 was obtained from a prostatic adenocarcinoma metastatic to bones and is a useful tool for modelling late stage cancer.[191] PC-3 also serves as a model for androgen independent PCa.[191] The PNT2C2 cell line, which is a subset of the PNT2 cell line, was obtained post mortem from a healthy prostate and immortalised with SV40200. This cell line is often used as a model of the normal prostate epithelium.[192] Previous studies have mainly focussed on gene expression and transcript profiles to identify differences in prostatic conditions.[193,194] These provide useful information on possible causes observed in protein expression changes and add to the knowledge base of prostatic conditions. However, protein products undergo multiple changes after transcription and translation and thus these findings may be less informative then a proteomic approach116.

A total of 3576 proteins were identified in this study and IPAs' extensive curated

database was utilised to compare the differences in pathways, upstream regulators and molecular and cellular functions between the cell lines (table 3.3 - 3.8 and figure 3.5 -3.7). In all cases the BPH-1, LNCaP and PC-3 cell lines were compared to the PNT2C2 reference cell line. The aim of these analyses was to determine if there were any differences in these cell lines, which could be exploited as potential biomarkers, especially as the current biomarker PSA does not discriminate between BPH and PCa.

It is perhaps unsurprising that IPA analysis yielded several common traits among the cell lines since they all essentially share the same origin (prostate epithelial cells), albeit from different disease states. This trend is clearly observed in the conserved nature of the pathways, upstream regulators, disease states, molecular and cellular functions and the assigned GO terms. GO term annotations showed extensive conservation in biological processing, cellular components and molecular function. This conservation therefore makes the identification of biomarkers unique to cancerous or benign conditions a challenge. IPA analysis did, however, identify specific pathways and upstream regulators upregulated in the LNCaP and PC-3 cell lines, but not the BPH-1 cell line that could potentially be exploited as possible avenues for biomarker discovery.

The EIF2 signalling, eIF4 and p70S6K signalling, tRNA charging, mTOR signalling and protein ubiquitination pathways were identified as upregulated in the three cell lines and are directly responsible for functions such as cell growth, proliferation, RNA post-transcriptional modification, protein synthesis and development. They act through various mechanisms such as translation initiation and control, protein recycling and pathway activation in response to stimuli and are critical in maintaining cell viability, cell growth potential, and proliferation. Interestingly, mitochondrial pathways and oxidative phosphorylation pathways were shown to be significantly upregulated in the LNCaP cell line compared to the BPH-1 and PC-3 cell lines. These pathways are responsible for maintaining ATP production and control of reactive oxygen species (ROS). Mitochondrial dysfunction and oxidative phosphorylation pathway disruptions may impact on effective ROS neutralisation, which subsequently results in DNA damage. This may result in and necrosis or apoptosis but can also result in mutant cellular proliferation. It is well known that changes in mitochondria function and oxidative phosphorylation have cancerous effects and in particular ROS have been implicated in carcinogenesis through oxidative stress and mitogenic signalling remodelling.[195] These modifications allow increased resistance to apoptosis in a oxidative phosphorylation disrupted low oxygen environment.[195] Notably, the remodelling of Adherens Junctions and epithelial Adherens Junctions pathways that are directly linked to cellular migration were significantly upregulated in both of the prostate cancer cell lines (LNCaP and PC-3) compared to the BPH-1 cell line. The Adherens Junctions and epithelial Adherens

Junctions are specialised sub-apical structures that control cell-cell adhesion and the remodelling of these junctions results in their dissociation and increases the migration potential of cells increasing the risk of metastasis. The unique nature of these pathways that are highly upregulated in the PCa related cell lines may provide an avenue for reliable biomarker discovery. Previous work on regulation of Adherens junctions and their effect on PCa previously been investigated by Ramteke et al. and Knights et al.[196,197] Ramteke et al. showed that proteins from exomes secreted by LNCaP and PC-3 cells under hypoxic conditions were associated with the remodelling of the epithelial Adherens junctions44. These secreted exomes and proteins were also shown to induce motility and invasiveness of other naïve LNCaP and PC-3 cells.[196] Knights et al. showed that the deregulation of Adherens junction genes is positively correlated with progression PCa and that Adherens junction components such as beta-catenin and E-cadherin, which were also identified in the LNCaP and PC-3 cell lines during this study, can act as prognostic and diagnostic biomarkers for PCa45. Exploitation of proteins preferentially involved in these pathways could therefore act potentially useful candidate biomarkers for PCa.

The core analysis function of IPA can be used to identify the possible disease states, which are a result of the identified upregulated pathways, upstream regulators and molecular and cellular functions. Cancer and tumour morphology were two of the conditions identified which explained the identified changes in all cell lines. A causal link can be seen between the upregulation of the EIF2 signalling, eIF4 and p70S6K signalling, tRNA charging, mTOR signalling and protein ubiquitination pathways and increased protein turnover, cell growth and proliferation which, are characteristic events associated with tumour and cancer progression.[198] The finding that cancer and tumour morphology were also identified as possible disease states for BPH-1 further emphasises the similarities between the PCa and BPH and thus the difficulty in identifying a biomarker which is unique to PCa and not BPH. Nonetheless, several known curated candidate biomarkers were identified (tables 3.14 and 3.15) in this study. What's more potentially novel biomarkers not in the IPA database were also identified (supplementary tables 5.1 - 5.3).

The identification of the known secreted proteins such as PSA and PSMA validated our approach in identifying new potential promising candidates.[153,193] Promising candidates included Armadillo repeat protein deleted in velo-cardio-facial syndrome (C9JJX6), Ankyrin-3 (Q12955), protein transport protein Sec23B (Fragment) (Q5F2F8) and Annexin A6 (P08133), which were all identified by IPA and were found in either LNCaP or PC-3 when considering only extracellular and membrane bound proteins. Protein LLP homolog (Protein LAPS18-like) (Q9BRT6) and Cell growth-regulating nucleolar protein (Q9NX58) were identified as common biomarkers for both LNCaP and PC-3 cell lines and are therefore the two of the most promising candidates. While Protein LLP homolog (Protein LAPS18-like) is

a cell permeable protein known to interact with transcriptional machinery such as RNA polymerase II and TATA binding protein aiding in translation and protein synthesis, this protein has not been well characterised to date and more work needs to be done to validate its potential as a biomarker for PCa. The Cell growth-regulating nucleolar protein was identified as membrane bound or extracellular protein in this study, a finding which contradicts previous studies which showed that it was preferentially localised to the nucleus.[199] Lishan et al. validated its subcellular location using indirect immunofluorescence staining which could explain why it was not detected as membrane bound or the extracellular component.[199] Analysis on the gene level has showed the presence of two putative DNA binding zinc finger domains which it uses to regulate cell growth through sequence specific binding and transcriptional regulation[199] Furthermore, Lishan et al. suggested that this protein may be involved in cell proliferation and tumourigenesis.[199]

While the identification of extracellular and membrane bound proteins from the whole proteome is useful as demonstrated above, an ideal biomarker should be detectable in serum or urine. The characterisation of the secretome provides a method which is more likely to lead to the discovery of such proteins and is essential for the validation of potential candidates.

## 4.5 Secretome analysis

Secretomics, which is the study of proteins secreted from cells during normal and altered cellular function has recently come to the forefront of biomarker discovery for its ability to rapidly identify a large set of proteins that are extracellular to cells and are therefore possibly non-invasive to obtain in a clinical setting.[172] The field of secretomics combined with advances in bottom up and quantitative proteomics has provided a powerful approach when identifying novel biomarkers.[172] The use of a bottom up approach has also been linked to incorrect identification of proteins and therefore must be taken into consideration when deciding on potential candidate biomarkers.[121]

As previously mentioned sample optimisation, collection and precipitation was completed on secretome samples. Unfortunately, due to technical difficulties only a preliminary analysis of preliminary secretome samples precipitated using ammonium sulphate was performed. These samples were not optimal due low recovery, which was later demonstrated and discussed above.

Nonetheless, distinct protein expression profiles were observed. The distribution of unique proteins across the cell lines (figure 3.13) shows that 3%, 24%, 9% and 4% of the identified proteins were unique to the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines, respectively. The subsequent selection of membrane bound and extracellular proteins further validated the changes in unique protein expression

and showed that 5%, 17%, 8% and 3% of the identified proteins were unique to the BPH-1, LNCaP, PC-3 and PNT2C2 cell lines, respectively. The selection of extracellular and membrane bound protein allowed the identification of several candidate biomarkers from the IPA database (table 3.17) and identified a number of novel candidate biomarkers for PCa (supplementary table 5.4 - 5.6). Interestingly, only a small overlap was observed between the proteome and secretome when considering membrane and extracellular proteins. This is perhaps not surprising as the analysis of proteome samples contain significantly more proteins which may mask the detection of candidate biomarkers. Furthermore, the proteome samples were washed prior to analysis, so any secreted proteins or loosely bound membrane proteins may have been removed. This finding therefore demonstrates the importance of utilising a secretome approach.

Ankyrin-3, Methylosome subunit pICln, Clathrin heavy chain 1, Ubiquitin-like protein ISG15, Moesin, Spectrin alpha chain non-erythrocytic 1, Spectrin beta chain non-erythrocytic 1 and DNA FLJ53769, highly similar to Homo sapiens nebulette (NEBL), transcript variant 2, mRNA were identified as potential biomarkers for PCa from IPA secretome analysis. Wei et al. previously showed that methylosome subunit pICln has been linked to PCa through its interaction with certain protein arginine methyltransferases (PRMT) such as PRMT5, which have been implicated in tumourigenesis.[200] Prescott et al. have shown that Clathrin heavy chain 1 was upregulated the presence of androgens in the LNCaP cell line and indicated that it may be involved in the exocytosis of androgen-regulated secretory proteins such as PSA.[201] Kiessling et al. have shown a positive link between ubiquitin-like protein ISG15 and PCa.[202] They showed that it was upregulated in LNCaP cells and further upregulated in patients who have undergone ablation therapy.[202] Chakraborty et al. have shown that Moesin can be positively correlated to PCa through its interaction with G protein-coupled receptor kinases (GRK) such GRK5 is known to increase the migration and invasion of PCa cells.[203] Khan et al. revealed that Spectrin alpha chain non-erythrocytic 1 was upregulated in PCa cells compared to benign tissue.[204] Zhixing et al. showed that there was a negative correlation between Spectrin beta chain non-erythrocytic 1 and activating transcription factor 4 (ATF4 which has previously been implicated in PCa.[205] Monitoring the loss of Spectrin beta chain non-erythrocytic 1 could allow the monitoring of PCa progression.[205] Wakamatsu et al. has identified the cDNA FLJ53769, highly similar to Homo sapiens nebulette (NEBL), transcript variant 2, mRNA but only at the transcript level for PCa.[206]

Ankyrin-3 is potentially the most promising candidate biomarker for PCa identified in this study as it was identified by IPA to be specific to the LNCaP proteome and secretome and was confirmed as LNCaP specific by our MS analysis with a high significance in both cases. Its main cellular functions have been

linked to cell motility, activation, proliferation, contact, and the maintenance of specialised membrane domains.[207] It has been identified at the protein level and implicated in multiple cancer types including PCa by The Human Protein Atlas group (www.proteinatlas.org). Its high upregulation, subcellular location and its unique association to the LNCaP cell line make it a potentially useful biomarker for the relatively early diagnosis of PCa.

## 4.6 Conclusion

The proteomic-based approach utilised in this study yielded multiple previously identified and new proteins that could be used as potential biomarkers for PCa. The analysis of the proteome allowed for a deeper understanding of the regulation of the cell based models in terms of preferential pathways, upstream regulators, and molecular and cellular processes. This study has shown that a secretomics based approach provides a simple and high throughput means of identifying multiple potential extracellular candidate biomarkers. As only the preliminary secretome data was available the possible candidate biomarkers would still need to be adequately validated before any further assumptions can be made on their potential. Furthermore, only proteins, which were found to be membrane bound or extracellular, were selected as potential candidate biomarkers. While this approach reduces the analysis space and increases the likelihood of identifying non-invasive biomarkers it relies on accurately assigned GO terms. Therefore by applying this type of filter several potential biomarkers may be lost. This preliminary secretome investigation validates a proof of concept for the identification of novel biomarkers for prostate cancer using a proteomic approach.

## 4.7 Future studies

Numerous future applications of this study are possible. Weaknesses in this study such as the presence of intracellular proteins need to be reduced further possibly by further optimisation of cell culture techniques or through the use of alternate cell culture methods such as 3D cell culture. The implications of identifying biomarkers in secretome cell culture studies is also only the first step in identifying novel tests for PCa, extensive upscaling needs to be done to confirm the proteins identified are present only in the cell lines observed and not false positive identifications from mass spectrometry. This study acts as a foundation for further studies utilising primary culture xenografts, patient sera and finally large-scale patient studies to confirm the specificity, sensitivity and efficacy of identified biomarkers.

# References


1. Stewart, B. W., Wild, C., International Agency for Research on Cancer. & World Health Organization. *World cancer report 2014* (International Agency for Research on Cancer, 2014).

2. Martin, T. A., Ye, L., Sanders, A. J., Lane, J. & Jiang, W. G. Cancer Invasion and Metastasis: Molecular and Cellular Perspective. *Journal of Proteomics* (2013).

3. Cooper, G. M. & Hausman, R. E. *The Cell : A Molecular Approach* (ASM Press, 2007), 4th edn edn. URL `https://www.ncbi.nlm.nih.gov/books/NBK9839/`.

4. Types of cancer — Cancer Research UK. URL `http://www.cancerresearchuk.org/about-cancer/what-is-cancer/how-cancer-starts/types-of-cancer`.

5. Eary, J. F. & Conrad, E. U. Imaging in sarcoma. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* **52**, 1903–13 (2011). URL `http://www.ncbi.nlm.nih.gov/pubmed/22052127`.

6. Cooper, D. A., Petts, V., Luckhurst, E., Biggs, J. C. & Penny, R. T and B cell populations in blood and lymph node in lymphoproliferative disease. *British journal of cancer* **31**, 550–8 (1975). URL `http://www.ncbi.nlm.nih.gov/pubmed/1080424http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2009434`.

7. Ulbright, T. M. Germ cell tumors of the gonads: a selective review emphasizing problems in differential diagnosis, newly appreciated, and controversial issues. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **18 Suppl 2**, S61–79 (2005). URL `http://www.ncbi.nlm.nih.gov/pubmed/15761467`.

8. Buckley, J. D. The aetiology of cancer in the very young. *The British journal of cancer. Supplement* **18**, S8–12 (1992). URL `http:`

//www.ncbi.nlm.nih.gov/pubmed/1323994http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2149654.

9. Henderson, B. E. & Feigelson, H. S. Hormonal carcinogenesis. *Carcinogenesis* **21**, 427–33 (2000). URL http://www.ncbi.nlm.nih.gov/pubmed/10688862.

10. Hanukoglu, I. Steroidogenic enzymes: structure, function, and role in regulation of steroid hormone biosynthesis. *The Journal of steroid biochemistry and molecular biology* **43**, 779–804 (1992). URL http://www.sciencedirect.com/science/article/pii/0960076092903075.

11. Hurst, D. R. & Welch, D. R. Metastasis suppressor genes at the interface between the environment and tumor cell growth. *International review of cell and molecular biology* **286**, 107–80 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/21199781http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3575029.

12. Fletcher, R. H. The diagnosis of colorectal cancer in patients with symptoms: finding a needle in a haystack. *BMC medicine* **7**, 18 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/19374737http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2672954.

13. Ellis, P. M. & Vandermeer, R. Delays in the diagnosis of lung cancer. *Journal of thoracic disease* **3**, 183–8 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/22263086http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3256519.

14. Countries, I. o. M. U. C. o. C. C. i. L., Middle-Income, Sloan, F. A. & Gelband, H. Cancer Causes and Risk Factors and the Elements of Cancer Control. *National Academy of Sciences* (2007).

15. Bhatt, M., Kant, S. & Bhaskar, R. Pulmonary tuberculosis as differential diagnosis of lung cancer. *South Asian journal of cancer* **1**, 36–42 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/24455507http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3876596.

16. Hamilton, W., Sharp, D. J., Peters, T. J. & Round, A. P. Clinical features of prostate cancer before diagnosis: a population-based, case-control study. *The British journal of general practice : the journal of the Royal College of General Practitioners* **56**, 756–62 (2006). URL http://www.ncbi.nlm.nih.gov/pubmed/17007705http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1920715.

17. Wardle, J., Robb, K., Vernon, S. & Waller, J. Screening for Prevention and Early Diagnosis of Cancer. *American Psychologist* (2015).

18. Sharma, S. Tumor markers in clinical practice: General principles and guidelines. *Indian journal of medical and paediatric oncology : official journal of Indian Society of Medical & Paediatric Oncology* **30**, 1–8 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/20668599http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2902207`.

19. Bray, F., Jemal, A., Grey, N., Ferlay, J. & Forman, D. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. *The lancet oncology* **13**, 790–801 (2012). URL `http://www.thelancet.com/journals/a/article/PIIS1470-2045{%}252812{%}252970211-5/fulltext`.

20. A, A. *Burden: mortality, morbidity and risk factors* (World Health Organisation (WHO), 2008).

21. Miranda, J. J., Kinra, S., Casas, J. P., Davey Smith, G. & Ebrahim, S. Non-communicable diseases in low- and middle-income countries: context, determinants and health policy. *Tropical medicine & international health : TM & IH* **13**, 1225–34 (2008). URL `http://www.ncbi.nlm.nih.gov/pubmed/18937743http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2687091`.

22. Worldwide cancer statistics — Cancer Research UK. URL `http://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer`.

23. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F.GLOBOCAN 2012 v1.0. Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 (2013). URL `http://globocan.iarc.fr`.

24. Lee, C. H., Akin-Olugbade, O. & Kirschenbaum, A. Overview of Prostate Anatomy, Histology, and Pathology. *Endocrinology and Metabolism Clinics of North America* **40**, 565–575 (2011).

25. Benninghoff, A. *Anatomie, Makroskopische Anatomie, Embryologie und Histologie des Menschen* (Urban und Schwarzenberg, München; Wien; Baltimore, 1993), 15 edn.

26. McNeal, J. E. Normal histology of the prostate. *The American journal of surgical pathology* **12**, 619–33 (1988). URL http://www.ncbi.nlm.nih.gov/pubmed/2456702.

27. Tisell, L. E. & Salander, H. The lobes of the human prostate. *Scandinavian journal of urology and nephrology* **9**, 185–91 (1975). URL http://www.ncbi.nlm.nih.gov/pubmed/1209174.

28. Martini, F. H., Timmons, M. J., & Tallitsch, R. B. *Human Anatomy* (Pearson Benjamin Cummings, San Francisco, 2012), 7th edn.

29. Selman, S. H. The McNeal prostate: a review. *Urology* **78**, 1224–8 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/21908026.

30. McNeal, J. E. Regional morphology and pathology of the prostate. *American journal of clinical pathology* **49**, 347–57 (1968). URL http://www.ncbi.nlm.nih.gov/pubmed/5645095.

31. Xia, S.-J., Xu, X.-X., Teng, J.-B., Xu, C.-X. & Tang, X.-D. Characteristic pattern of human prostatic growth with age. *Asian journal of andrology* **4**, 269–71 (2002). URL http://www.ncbi.nlm.nih.gov/pubmed/12508127.

32. Vargas, H. A. *et al.* Normal central zone of the prostate and central zone involvement by prostate cancer: clinical and MR imaging implications. *Radiology* **262**, 894–902 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22357889http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3285222.

33. Cohen, R. J. *et al.* Central Zone Carcinoma of the Prostate Gland: A Distinct Tumor Type With Poor Prognostic Features. *The Journal of Urology* **179**, 1762–1767 (2008). URL http://linkinghub.elsevier.com/retrieve/pii/S0022534708000207.

34. Oh, W. K., Hurwitz, M., D'Amico, A. V., Richie, J. P. & Kantoff, P. W. *Biology of Prostate Cancer* (BC Decker, 2003).

35. Villeirs, G. M., Verstraete, K. L., De Neve, W. J. & De Meerleer, G. O. Magnetic resonance imaging anatomy of the prostate and periprostatic area: a guide for radiotherapists. *Radiotherapy and Oncology* (2005).

36. Ayala, A. G., Ro, J. Y., Babaian, R., Troncoso, P. & Grignon, D. J. The prostatic capsule: does it exist? Its importance in the staging and treatment of prostatic carcinoma. *The American journal of surgical pathology* **13**, 21–7 (1989). URL http://www.ncbi.nlm.nih.gov/pubmed/2909195.

37. Denmeade, S. R. & Isaacs, J. T. *Cellular Organization of the Normal Prostate* (BC Decker, 2003). URL https://www.ncbi.nlm.nih.gov/books/NBK14022/.

38. Online, I. H. How does the prostate work? *IQWiG (Institute for Quality and Efficiency in Health Care)* (2016).

39. Talwar, G. P. *Textbook of biochemistry, biotechnology, allied and molecular medicine.* (Prentice-Hall Of India, 2015).

40. LeBeau, A. M., Kostova, M., Craik, C. S. & Denmeade, S. R. Prostate-specific antigen: an overlooked candidate for the targeted treatment and selective imaging of prostate cancer. *Biological chemistry* **391**, 333–43 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20180648http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3454521.

41. de Lamirande, E. Semenogelin, the Main Protein of the Human Semen Coagulum, Regulates Sperm Function. *Seminars in Thrombosis and Hemostasis* **33**, 060–068 (2007). URL http://www.ncbi.nlm.nih.gov/pubmed/17253191http://www.thieme-connect.de/DOI/DOI?10.1055/s-2006-958463.

42. Arnold, J. T. & Isaacs, J. T. Mechanisms involved in the progression of androgen-independent prostate cancers: it is not only the cancer cell's fault. *Endocrine-related cancer* **9**, 61–73 (2002). URL http://www.ncbi.nlm.nih.gov/pubmed/11914183http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4124629.

43. Carlsson, L. Perspectives on the Biological Role of Human Prostasomes. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* (2001).

44. Eikenberry, S. E., Nagy, J. D. & Kuang, Y. The evolutionary impact of androgen levels on prostate cancer in a multi-scale mathematical model. *Biology direct* **5**, 24 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20406442http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2885348.

45. Lipsett, M. B. *et al.* Studies on Leydig cell physiology and pathology: secretion and metabolism of testosterone. *Recent progress in hormone research* **22**, 245–81 (1966). URL http://www.ncbi.nlm.nih.gov/pubmed/5334626.

46. Bain, J. The many faces of testosterone. *Clinical interventions in aging* **2**, 567–76 (2007). URL http://www.ncbi.nlm.nih.gov/pubmed/18225457http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2686330.

47. Askew, E. B., Gampe, R. T., Stanley, T. B., Faggart, J. L. & Wilson, E. M. Modulation of androgen receptor activation function 2 by testosterone and dihydrotestosterone. *The Journal of biological chemistry* **282**, 25801–16 (2007). URL http://www.ncbi.nlm.nih.gov/pubmed/17591767http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4075031.

48. Edwards, J., Bartlett, J. M. S. & Edwards, J. The Androgen Receptor and Signal Transduction Pathways in. *British Journal of Urology International* **95**, 1320–1326 (2005). URL http://eprints.gla.ac.uk/7799/http://eprints.gla.ac.uk.

49. Wright, A. S., Douglas, R. C., Thomas, L. N., Lazier, C. B. & Rittmaster, R. S. Androgen-induced regrowth in the castrated rat ventral prostate: role of 5alpha-reductase. *Endocrinology* **140**, 4509–15 (1999). URL http://www.ncbi.nlm.nih.gov/pubmed/10499505.

50. Hudak, S. J., Hernandez, J. & Thompson, I. M. Role of 5 alpha-reductase inhibitors in the management of prostate cancer. *Clinical interventions in aging* **1**, 425–31 (2006). URL http://www.ncbi.nlm.nih.gov/pubmed/18046919http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2699636.

51. Steers, W. D. 5alpha-reductase activity in the prostate. *Urology* **58**, 17–24; discussion 24 (2001). URL http://www.ncbi.nlm.nih.gov/pubmed/11750244.

52. Zhu, Y.-S. & Sun, G.-H. 5$\alpha$-Reductase Isozymes in the Prostate. *Journal of medical sciences (Taipei, Taiwan)* **25**, 1–12 (2005). URL http://www.ncbi.nlm.nih.gov/pubmed/18483578http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2386416.

53. Ramakrishnan, K. & Salinas, R. C. Prostatitis: acute and chronic. *Primary care* **37**, 547–63, viii–ix (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20705198.

54. Dhingra, N. & Bhagwat, D. Benign prostatic hyperplasia: An overview of existing treatment. *Indian journal of pharmacology* **43**, 6–12 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/21455413http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3062123.

55. Roehrborn, C. G. Benign prostatic hyperplasia: an overview. *Reviews in urology* **7 Suppl 9**, S3–S14 (2005). URL http://www.ncbi.nlm.nih.gov/pubmed/16985902http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1477638.

56. Marihart, S., Harik, M. & Djavan, B. Dutasteride: a review of current data on a novel dual inhibitor of 5alpha reductase. *Reviews in urology* **7**, 203–10 (2005). URL http://www.ncbi.nlm.nih.gov/pubmed/16985831http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1550785.

57. Napalkov, P., Maisonneuve, P. & Boyle, P. Worldwide patterns of prevalence and mortality from benign prostatic hyperplasia. *Urology* **46**, 41–6 (1995). URL http://www.ncbi.nlm.nih.gov/pubmed/7544516.

58. Duncan, M. E. & Goldacre, M. J. Mortality trends for benign prostatic hyperplasia and prostate cancer in English populations 1979-2006. *BJU International* **107**, 40–45 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/20590542http://doi.wiley.com/10.1111/j.1464-410X.2010.09487.x.

59. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, C. K. e. SEER Cancer Statistics Review, 1975-2013 (2015). URL http://seer.cancer.gov/csr/1975{_}2013/.

60. Pakzad, R., Mohammadian-Hafshejani, A., Ghoncheh, M., Pakzad, I. & Salehiniya, H. The incidence and mortality of prostate cancer and its relationship with development in Asia. *Prostate international* **3**, 135–40 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/26779461http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4685206.

61. Researched *et al.* Cancer Association of South Africa (CANSA) Fact Sheet and Position Statement on Prostate Cancer. *Health StudiesBA Social Work* (2016).

62. Cuzick, J. *et al.* Prevention and early detection of prostate cancer. *The Lancet. Oncology* **15**, e484–92 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/25281467http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4203149.

63. PDQ Cancer Genetics Editorial Board, P. C. G. E. *Genetics of Prostate Cancer (PDQ®): Health Professional Version* (National Cancer Institute (US), 2002). URL http://www.ncbi.nlm.nih.gov/pubmed/26389227.

64. Key, T. J. *et al.* Diet, nutrition and the prevention of cancer. *Public health nutrition* **7**, 187–200 (2004). URL http://www.ncbi.nlm.nih.gov/pubmed/14972060.

65. Castro, E. & Eeles, R. The role of BRCA1 and BRCA2 in prostate cancer. *Asian journal of andrology* **14**, 409–14 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22522501http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3720154.

66. Abate-Shen, C. & Shen, M. M. Molecular genetics of prostate cancer. *Genes & development* **14**, 2410–34 (2000). URL http://www.ncbi.nlm.nih.gov/pubmed/11018010.

67. Lee, D. K. *et al.* Progression of prostate cancer despite an extremely low serum level of prostate-specific antigen. *Korean journal of urology* **51**, 358–61 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20495701http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2873892.

68. Bubendorf, L. *et al.* Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Human pathology* **31**, 578–83 (2000). URL http://www.ncbi.nlm.nih.gov/pubmed/10836297.

69. Wong, S. Y. & Hynes, R. O. Lymphatic or hematogenous dissemination: how does a metastatic tumor cell decide? *Cell cycle (Georgetown, Tex.)* **5**, 812–7 (2006). URL http://www.ncbi.nlm.nih.gov/pubmed/16627996http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1459485.

70. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013). URL http://www.nature.com/doifinder/10.1038/nature12477.

71. Loeb, K. R. & Loeb, L. A. Significance of multiple mutations in cancer. *Carcinogenesis* **21**, 379–85 (2000). URL http://www.ncbi.nlm.nih.gov/pubmed/10688858.

72. Shtivelman, E., Beer, T. M. & Evans, C. P. Molecular pathways and targets in prostate cancer. *Oncotarget* **5**, 7217–59 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/25277175http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4202120.

73. Cavallo, F., De Giovanni, C., Nanni, P., Forni, G. & Lollini, P. L. 2011: The immune hallmarks of cancer. In *Cancer Immunology, Immunotherapy*, vol. 60, 319–326 (2011). URL http:

//www.ncbi.nlm.nih.gov/pubmed/23209317http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3526364. 0208024.

74. Wong, R. S. Y. Apoptosis in cancer: from pathogenesis to treatment. *Journal of experimental & clinical cancer research : CR* **30**, 87 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/21943236http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3197541.

75. Robinson, D. *et al.* Integrative clinical genomics of advanced prostate cancer. *Cell* **161**, 1215–1228 (2015). 15334406.

76. McCain, J. The MAPK (ERK) Pathway: Investigational Combinations for the Treatment Of BRAF-Mutated Metastatic Melanoma. *P & T : a peer-reviewed journal for formulary management* **38**, 96–108 (2013). URL http://www.ncbi.nlm.nih.gov/pubmed/23599677http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3628180.

77. Ilmer, M. *et al.* RSPO2 Enhances Canonical Wnt Signaling to Confer Stemness-Associated Traits to Susceptible Pancreatic Cancer Cells. *Cancer Research* **75**, 1883–1896 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/25769727http://cancerres.aacrjournals.org/cgi/doi/10.1158/0008-5472.CAN-14-1327.

78. Owonikoko, T. K. & Khuri, F. R. Targeting the PI3K/AKT/mTOR pathway: biomarkers of success and tribulation. *American Society of Clinical Oncology educational book. American Society of Clinical Oncology. Meeting* (2013). URL http://www.ncbi.nlm.nih.gov/pubmed/23714559http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3821994.

79. Saramaki, O. & Visakorpi, T. Chromosomal aberrations in prostate cancer. *Frontiers in bioscience : a journal and virtual library* **12**, 3287–301 (2007). URL http://www.ncbi.nlm.nih.gov/pubmed/17485299.

80. Yardy, G. W. & Brewster, S. F. Wnt signalling and prostate cancer. *Prostate Cancer and Prostatic Diseases* **8**, 119–126 (2005). URL http://www.nature.com/doifinder/10.1038/sj.pcan.4500794.

81. Zhao, Y., Tindall, D. J. & Huang, H. Modulation of Androgen Receptor by FOXA1 and FOXO1 Factors in Prostate Cancer. *International Journal of Biological Sciences* **10**, 614–619 (2014). URL http://www.ijbs.com/v10p0614.htm.

82. Xiao, G.-Q. *et al.* Loss of PLZF Expression in Prostate Cancer by Immunohistochemistry Correlates with Tumor Aggressiveness and Metastasis. *PLOS*

*ONE* **10**, e0121318 (2015). URL http://dx.plos.org/10.1371/journal.pone.0121318.

83. Boormans, J. L. *et al.* E17K substitution in AKT1 in prostate cancer. *British journal of cancer* **102**, 1491–4 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20407443http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2869172.

84. Edlind, M. P. & Hsieh, A. C. PI3K-AKT-mTOR signaling in prostate cancer progression and androgen deprivation therapy resistance. *Asian journal of andrology* **16**, 378–86 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/24759575http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4023363.

85. Sharma, A. *et al.* The retinoblastoma tumor suppressor controls androgen signaling and human prostate cancer progression. *The Journal of clinical investigation* **120**, 4478–92 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/21099110http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2993601.

86. Gudmundsdottir, K. & Ashworth, A. The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. *Oncogene* **25**, 5864–74 (2006). URL http://www.ncbi.nlm.nih.gov/pubmed/16998501.

87. Hiom, S. C. Diagnosing cancer earlier: reviewing the evidence for improving cancer survival. *British journal of cancer* S1–5 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/25734391http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4385969.

88. Yokota, J. Tumor progression and metastasis. *Carcinogenesis* **21**, 497–503 (2000). URL http://www.ncbi.nlm.nih.gov/pubmed/10688870.

89. Kaur, J. & Mohanti, B. K. Transition from curative to palliative care in cancer. *Indian journal of palliative care* **17**, 1–5 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/21633614http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3098537.

90. Takes, R. P. *et al.* Future of the TNM classification and staging system in head and neck cancer. *Head & Neck* **32**, 1693–1711 (2010). URL http://doi.wiley.com/10.1002/hed.21361.

91. Harrington, S. E. & Smith, T. J. The role of chemotherapy at the end of life: "when is enough, enough?". *JAMA* **299**, 2667–78

(2008). URL http://www.ncbi.nlm.nih.gov/pubmed/18544726http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3099412.

92. Ambrus, J. L., Ambrus, C. M., Mink, I. B. & Pickren, J. W. Causes of death in cancer patients. *Journal of medicine* **6**, 61–4 (1975). URL http://www.ncbi.nlm.nih.gov/pubmed/1056415.

93. Montazeri, A. Quality of life data as prognostic indicators of survival in cancer patients: an overview of the literature from 1982 to 2008. *Health and quality of life outcomes* **7**, 102 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/20030832http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2805623.

94. Michaelson, M. D. *et al.* Management of complications of prostate cancer treatment. *CA: a cancer journal for clinicians* **58**, 196–213 (2008). URL http://www.ncbi.nlm.nih.gov/pubmed/18502900http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2900775.

95. Efstathiou, J. A., Shipley, W. U., Zietman, A. L. & Smith, M. R. Hormonal therapies: ADT for prostate cancer: true love or heartbreak? *Nature reviews. Clinical oncology* **7**, 130–2 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20190794http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2900769.

96. Challapalli, A., Jones, E., Harvey, C., Hellawell, G. O. & Mangar, S. A. High dose rate prostate brachytherapy: an overview of the rationale, experience and emerging applications in the treatment of prostate cancer. *The British journal of radiology* S18–27 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/23118099http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3746400.

97. Kolinsky, M., Rescigno, P. & de Bono, J. S. Chemical or Surgical CastrationIs This Still an Important Question? *JAMA Oncology* **2**, 437 (2016). URL http://oncology.jamanetwork.com/article.aspx?doi=10.1001/jamaoncol.2015.4918.

98. Potosky, A. L. *et al.* Effectiveness of Primary Androgen-Deprivation Therapy for Clinically Localized Prostate Cancer. *Journal of Clinical Oncology* **32**, 1324–1330 (2014). URL http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2013.52.5782.

99. Peehl, D. M. Basic science of hormonal therapy for prostate cancer. *Reviews in urology* **3 Suppl 3**, S15–22 (2001). URL http:

//www.ncbi.nlm.nih.gov/pubmed/16986004http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1476082.

100. Hotte, S. J. & Saad, F. Current management of castrate-resistant prostate cancer. *Current oncology (Toronto, Ont.)* **17 Suppl 2**, S72–9 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20882137http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2935714.

101. Perlmutter, M. A. & Lepor, H. Androgen deprivation therapy in the treatment of advanced prostate cancer. *Reviews in urology* **9 Suppl 1**, S3–8 (2007). URL http://www.ncbi.nlm.nih.gov/pubmed/17387371http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1831539.

102. Strope, S. A. & Andriole, G. L. Prostate cancer screening: current status and future perspectives. *Nature reviews. Urology* **7**, 487–93 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20818326.

103. Borley, N. & Feneley, M. R. Prostate cancer: diagnosis and staging. *Asian journal of andrology* **11**, 74–80 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/19050692http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3735216.

104. Loeb, S. & Catalona, W. J. The Prostate Health Index: a new test for the detection of prostate cancer. *Therapeutic advances in urology* **6**, 74–7 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/24688603http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3943368.

105. Charatan, F. Screening for prostate cancer may not reduce mortality. *BMJ : British Medical Journal* **332**, 72 (2006).

106. Madu, C. O. & Lu, Y. Novel diagnostic biomarkers for prostate cancer. *Journal of Cancer* **1**, 150–77 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20975847http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2962426.

107. Shariat, S. F. & Roehrborn, C. G. Using biopsy to detect prostate cancer. *Reviews in urology* **10**, 262–80 (2008). URL http://www.ncbi.nlm.nih.gov/pubmed/19145270http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2615104.

108. Madej, A., Wilkosz, J., Różański, W. & Lipiński, M. Complication rates after prostate biopsy according to the number of sampled cores. *Central European journal of urology* **65**, 116–8

(2012). URL http://www.ncbi.nlm.nih.gov/pubmed/24578945http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3921797.

109. Mishra, A. & Verma, M. Cancer biomarkers: are we ready for the prime time? *Cancers* **2**, 190–208 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/24281040http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3827599.

110. Strimbu, K. & Tavel, J. A. What are biomarkers? *Current opinion in HIV and AIDS* **5**, 463–6 (2010). URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3078627{&}tool=pmcentrez{&}rendertype=abstract.

111. Kulasingam, V. & Diamandis, E. P. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nature Clinical Practice Oncology* **5**, 588–599 (2008). URL http://www.nature.com/doifinder/10.1038/ncponc1187.

112. Saah, A. J. & Hoover, D. R. "Sensitivity" and "specificity" reconsidered: the meaning of these terms in analytical and diagnostic settings. *Annals of internal medicine* **126**, 91–4 (1997). URL http://www.ncbi.nlm.nih.gov/pubmed/8992938.

113. Rao, A. R., Motiwala, H. G. & Karim, O. M. A. The discovery of prostate-specific antigen. *BJU international* **101**, 5–10 (2008). URL http://www.ncbi.nlm.nih.gov/pubmed/17760888.

114. Kyung, Y.-S., Lee, H.-C. & Kim, H.-J. Changes in serum prostate-specific antigen after treatment with antibiotics in patients with lower urinary tract symptoms/benign prostatic hyperplasia with prostatitis. *International neurourology journal* **14**, 100–4 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/21120219http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2989471.

115. Serefoglu, E. C. *et al.* How reliable is 12-core prostate biopsy procedure in the detection of prostate cancer? *Canadian Urological Association journal = Journal de l'Association des urologues du Canada* **7**, E293–8 (2013). URL http://www.ncbi.nlm.nih.gov/pubmed/22398204http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3668408.

116. Horgan, R. P., Kenny, L. C., Mrcpi, M. & Phd, M. SAC review 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics Clinical Research Fellow and Specialist Registrar in Obstetrics and Gynaecology. *SAC*

*review SAC review The Obstetrician & Gynaecologist* **1313** (2011). URL http://onlinetog.org.

117. Wilson, K. E. *et al.* Functional genomics and proteomics: application in neurosciences. *J Neurol Neurosurg Psychiatry* **75**, 529–538 (2004).

118. Johann, D. J. *et al.* Clinical Proteomics and Biomarker Discovery. *Annals of the New York Academy of Sciences* **1022**, 295–305 (2004). URL http://doi.wiley.com/10.1196/annals.1318.045.

119. Karagiannis, G. S. *et al.* Cancer secretomics reveal pathophysiological pathways in cancer molecular oncology. *Molecular oncology* **4**, 496–510 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20934395.

120. Chait, B. T. Mass Spectrometry: Bottom-Up or Top-Down? *Science* **314** (2006).

121. Han, X., Aslanian, A., Yates, J. R. & III. Mass spectrometry for proteomics. *Current opinion in chemical biology* **12**, 483–90 (2008). URL http://www.ncbi.nlm.nih.gov/pubmed/18718552http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2642903.

122. Zhang, Y. *et al.* Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews* **113**, 2343–94 (2013). URL http://www.ncbi.nlm.nih.gov/pubmed/23438204http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3751594.

123. Ho, C. S. *et al.* Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical biochemist. Reviews* **24**, 3–12 (2003). URL http://www.ncbi.nlm.nih.gov/pubmed/18568044http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1853331.

124. De Laeter, J. R. The role of off-line mass spectrometry in nuclear fission. *Mass Spectrometry Reviews* **15**, 261–281 (1996). URL http://doi.wiley.com/10.1002/{%}28SICI{%}291098-2787{%}281996{%}2915{%}3A4{%}3C261{%}3A{%}3AAID-MAS3{%}3E3.0.CO{%}3B2-G.

125. Sidoli, S., Fujiwara, R. & Garcia, B. A. Multiplexed data independent acquisition (MSX-DIA) applied by high resolution mass spectrometry improves quantification quality for the analysis of histone peptides. *Proteomics* **16**, 2095–105 (2016). URL http://www.ncbi.nlm.nih.gov/pubmed/27193262http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5104670.

126. Canterbury, J. *et al.* Improvements for High Resolution Analysis on a Modified Tribrid Mass Spectrometer Improvements for high resolution analysis on a modified tribrid mass spectrometer. *Thermo Fisher Scientifics* (2015).

127. Parasuraman, S., Balamurugan, S., Muralidharan, S., Jayaraj Kumar, K. & Vijayan, V. An Overview of Liquid Chromatography-Mass Spectroscopy Instrumentation. *Pharmaceutical Methods* **5** (2014).

128. Frank, A. M. Algorithms for tandem mass spectrometry-based proteomics (2008).

129. Chernushevich, I. V., Loboda, A. V. & Thomson, B. A. An introduction to quadrupoletime-of-flight mass spectrometry. *Journal Of Mass Spectrometry J. Mass Spectrom* **36**, 849–865 (2001).

130. Hu, Q. *et al.* The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry* **40**, 430–443 (2005). URL `http://doi.wiley.com/10.1002/jms.856`.

131. Richard, H., Perry, R., Cooks, G. & Noll, R. J. Mass Spectrometry Fundamental LC-MS Orbitrap Mass Analyzers. *Wiley InterScience* (2008).

132. Catherman, A. D., Skinner, O. S. & Kelleher, N. L. Top Down proteomics: facts and perspectives. *Biochemical and biophysical research communications* **445**, 683–93 (2014). URL `http://www.ncbi.nlm.nih.gov/pubmed/24556311http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4103433`.

133. Gundry, R. L. *et al.* Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow. *Current protocols in molecular biology* **Chapter 10**, Unit10.25 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/19816929http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2905857`.

134. Zhang, G. *et al.* Protein quantitation using mass spectrometry. *Methods in molecular biology (Clifton, N.J.)* **673**, 211–22 (2010). URL `http://www.ncbi.nlm.nih.gov/pubmed/20835801http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3758905`.

135. Amunugama, R., Jones, R., Ford, M. & Allen, D. Bottom-Up Mass Spectrometry-Based Proteomics as an Investigative Analytical Tool for Discovery and Quantification of Proteins in Biological Samples. *Advances in wound care* **2**, 549–557 (2013).

URL http://www.ncbi.nlm.nih.gov/pubmed/24761338http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3842888.

136. Tu, C. *et al.* Optimization of Search Engines and Postprocessing Approaches to Maximize Peptide and Protein Identification for High-Resolution Mass Data. *Journal of Proteome Research* **14**, 4662–4673 (2015). URL http://pubs.acs.org/doi/10.1021/acs.jproteome.5b00536.

137. Deutsch, E. W. File formats commonly used in mass spectrometry proteomics. *Molecular & cellular proteomics : MCP* **11**, 1612–21 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22956731http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3518119.

138. Thiede, B. *et al.* Peptide mass fingerprinting. *Methods* **35**, 237–247 (2005).

139. Ma, B. Novor: real-time peptide de novo sequencing software. *Journal of the American Society for Mass Spectrometry* **26**, 1885–94 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/26122521http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4604512.

140. Horai, H. *et al.* MassBank: a public repository for sharing mass spectral data for life sciences. *Journal of Mass Spectrometry* **45**, 703–714 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20623627http://doi.wiley.com/10.1002/jms.1777.

141. Sugimoto, M., Kawakami, M., Robert, M., Soga, T. & Tomita, M. Bioinformatics Tools for Mass Spectroscopy-Based Metabolomic Data Processing and Analysis. *Current bioinformatics* **7**, 96–108 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22438836http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3299976.

142. Zhou, T. *et al.* MaxReport: An Enhanced Proteomic Result Reporting Tool for MaxQuant. *PloS one* **11**, e0152067 (2016). URL http://www.ncbi.nlm.nih.gov/pubmed/27003708http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4803341.

143. Vaudel, M. *et al.* PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nature Biotechnology* **33**, 22–24 (2015). URL http://www.nature.com/doifinder/10.1038/nbt.3109.

144. Gonzalez-Galarza, F. F., Qi, D., Fan, J., Bessant, C. & Jones, A. R. A tutorial for software development in quantitative proteomics using PSI standard formats. *Biochimica et biophysica acta* **1844**, 88–97

(2014). URL http://www.ncbi.nlm.nih.gov/pubmed/23584085http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4008935.

145. Gonzalez-Galarza, F. F. *et al.* A critical appraisal of techniques, software packages, and standards for quantitative proteomic analysis. *Omics : a journal of integrative biology* **16**, 431–42 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22804616http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3437040.

146. Horoszewicz, J. S. *et al.* LNCaP model of human prostatic carcinoma. *Cancer research* **43**, 1809–18 (1983). URL http://www.ncbi.nlm.nih.gov/pubmed/6831420.

147. Sanchez-Sweatman, O. H., Orr, F. W. & Singh, G. Human metastatic prostate PC3 cell lines degrade bone using matrix metalloproteinases. *Invasion & metastasis* **18**, 297–305 (1998). URL http://www.ncbi.nlm.nih.gov/pubmed/10729774.

148. Hayward, S. W. *et al.* Establishment and characterization of an immortalized but non-transformed human prostate epithelial cell line: BPH-1. *In vitro cellular & developmental biology. Animal* **31**, 14–24 (1995). URL http://www.ncbi.nlm.nih.gov/pubmed/7535634.

149. Pickard, M. R., Darling, D., Farzaneh, F. & Williams, G. T. Preparation and Characterization of Prostate Cell Lines for Functional Cloning Studies to Identify Regulators of Apoptosis. *Journal of Andrology* **30**, 248–258 (2009). URL http://doi.wiley.com/10.2164/jandrol.108.005686.

150. Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *ANALYTICAL BIOCHEMISTRY* **72**, 248–254 (1976).

151. Zor, T. & Selinger, Z. Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies. *Analytical biochemistry* **236**, 302–8 (1996). URL http://www.ncbi.nlm.nih.gov/pubmed/8660509.

152. Cytotoxicity Detection Kit (LDH) Protocol & Troubleshooting — Sigma-Aldrich. URL http://www.sigmaaldrich.com/technical-documents/protocols/biology/roche/cytotoxicity-detection-kit-ldh.html.

153. Sardana, G., Jung, K., Stephan, C. & Diamandis, E. P. Proteomic Analysis of Conditioned Media from the PC3, LNCaP, and 22Rv1 Prostate Cancer Cell Lines: Discovery and Validation of Candidate Prostate Cancer Biomarkers.

*Journal of Proteome Research* **7**, 3329–3338 (2008). URL http://pubs.acs.org/doi/abs/10.1021/pr8003216.

154. Merrell, K. *et al.* Analysis of low-abundance, low-molecular-weight serum proteins using mass spectrometry. *Journal of biomolecular techniques : JBT* **15**, 238–48 (2004). URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2291707{&}tool=pmcentrez{&}rendertype=abstract.

155. Feist, P. & Hummon, A. B. Proteomic challenges: sample preparation techniques for microgram-quantity protein analysis from biological samples. *International journal of molecular sciences* **16**, 3537–63 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/25664860http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4346912.

156. Peyrl, A., Krapfenbauer, K., Slavc, I., Strobel, T. & Lubec, G. Proteomic characterization of the human cortical neuronal cell line HCN-2. *Journal of chemical neuroanatomy* **26**, 171–8 (2003). URL http://www.ncbi.nlm.nih.gov/pubmed/14615026.

157. Stastna, M. & Van Eyk, J. E. Investigating the Secretome. *Circulation: Cardiovascular Genetics* **5** (2012).

158. Taylor, C. F. *et al.* Guidelines for reporting the use of mass spectrometry in proteomics. *Nature Biotechnology* **26**, 860–861 (2008). URL http://www.nature.com/doifinder/10.1038/nbt0808-860.

159. Martínez-Bartolomé, S. *et al.* Guidelines for reporting quantitative mass spectrometry based experiments in proteomics. *Journal of Proteomics* **95**, 84–88 (2013). URL http://linkinghub.elsevier.com/retrieve/pii/S1874391913001024.

160. Binz, P.-A. *et al.* Guidelines for reporting the use of mass spectrometry informatics in proteomics. *Nature Biotechnology* **26**, 862–862 (2008). URL http://www.nature.com/doifinder/10.1038/nbt0808-862.

161. Papanikolaou, E. *et al.* Cell Cycle Status of CD34+ Hemopoietic Stem Cells Determines Lentiviral Integration in Actively Transcribed and Development-related Genes. *Molecular Therapy* **23**, 683–696 (2015). URL http://linkinghub.elsevier.com/retrieve/pii/S1525001616300880.

162. Attwood, T. K., Pettifer, S. R. S. R. & Thorne, D. *Bioinformatics challenges at the interface of biology and computer science : mind the gap* (Wiley-Blackwell, 2016).

163. Bhatia, V. N., Perlman, D. H., Costello, C. E. & McComb, M. E. Software tool for researching annotations of proteins: open-source protein annotation software with data visualization. *Analytical chemistry* **81**, 9819–23 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/19839595http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2787672.

164. Planque, C. *et al.* Identification of Five Candidate Lung Cancer Biomarkers by Proteomics Analysis of Conditioned Media of Four Lung Cancer Cell Lines. *Molecular & Cellular Proteomics* **8**, 2746–2758 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/19776420http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2816016http://www.mcponline.org/cgi/doi/10.1074/mcp.M900134-MCP200.

165. Prensner, J. R., Rubin, M. A., Wei, J. T. & Chinnaiyan, A. M. Beyond PSA: the next generation of prostate cancer biomarkers. *Science translational medicine* **4**, 127rv3 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22461644http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3799996.

166. Slatkoff, S., Gamboa, S., Zolotor, A. J., Mounsey, A. L. & Jones, K. PURLs: PSA testing: when it's useful, when it's not. *The Journal of family practice* **60**, 357–60 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/21647471http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3183963.

167. Saini, S. PSA and beyond: alternative prostate cancer biomarkers. *Cellular oncology (Dordrecht)* **39**, 97–106 (2016). URL http://www.ncbi.nlm.nih.gov/pubmed/26790878http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4821699.

168. Ablin, R. J. & Haythorn, M. R. Screening for prostate cancer: Controversy? What controversy? *Current oncology (Toronto, Ont.)* **16**, 1–2 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/19526077http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2695709.

169. Lin, K., Croswell, J. M., Koenig, H., Lam, C. & Maltz, A. *Prostate-Specific Antigen-Based Screening for Prostate Cancer* (Agency for Healthcare Research and Quality (US), 2011). URL http://www.ncbi.nlm.nih.gov/pubmed/22171385.

170. Crawford, E. D., Ventii, K. & Shore, N. D. New biomarkers in prostate cancer. *Oncology (Williston Park, N.Y.)* **28**, 135–42 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/24701701.

171. Hu, B. *et al.* Discovering cancer biomarkers from clinical samples by protein microarrays. *PROTEOMICS - Clinical Applications* **9**, 98–110 (2015). URL http://doi.wiley.com/10.1002/prca.201400094.

172. Xue, H., Lu, B. & Lai, M. The cancer secretome: a reservoir of biomarkers. *Journal of translational medicine* **6**, 52 (2008). URL http://www.ncbi.nlm.nih.gov/pubmed/18796163http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2562990.

173. Makridakis, M. & Vlahou, A. Secretome proteomics for discovery of cancer biomarkers. *Journal of Proteomics* **73**, 2291–2305 (2010).

174. Kulasingam, V. & Diamandis, E. P. Proteomics analysis of conditioned media from three breast cancer cell lines: a mine for biomarkers and therapeutic targets. *Molecular & cellular proteomics : MCP* **6**, 1997–2011 (2007). URL http://www.ncbi.nlm.nih.gov/pubmed/17656355.

175. Gunawardana, C. G. *et al.* Comprehensive Analysis of Conditioned Media from Ovarian Cancer Cell Lines Identifies Novel Candidate Markers of Epithelial Ovarian Cancer. *Journal of Proteome Research* **8**, 4705–4713 (2009). URL http://pubs.acs.org/doi/abs/10.1021/pr900411g.

176. Peh, G. S. L. *et al.* Optimization of human corneal endothelial cell culture: density dependency of successful cultures in vitro. *BMC research notes* **6**, 176 (2013). URL http://www.biomedcentral.com/1756-0500/6/176.

177. Chevallet, M., Diemer, H., Van Dorssealer, A., Villiers, C. & Rabilloud, T. Toward a better analysis of secreted proteins: the example of the myeloid cells secretome. *Proteomics* **7**, 1757–70 (2007). URL http://www.ncbi.nlm.nih.gov/pubmed/17464941http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2386146.

178. Wingfield, P. Protein precipitation using ammonium sulfate. *Current protocols in protein science* **Appendix 3**, Appendix 3F (2001). URL http://www.ncbi.nlm.nih.gov/pubmed/18429073http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4817497.

179. Bonner, P. L. R. *Protein purification* (Taylor & Francis, 2007).

180. Grigera, J. R. & McCarthy, A. N. The behavior of the hydrophobic effect under pressure and protein denaturation. *Biophysical journal* **98**, 1626–31 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/20409483http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2856145.

181. Scott, J. E. Phosphotungstate: a "universal" (nonspecific) precipitant for polar polymers in acid solution. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **19**, 689–91 (1971). URL http://www.ncbi.nlm.nih.gov/pubmed/5121870.

182. Chatterjea, M. N. & Shinde, R. *Textbook of medical biochemistry* (Jaypee Brothers Medical Publications (P) Ltd, 2012).

183. Cutler, P. *Protein purification protocols.* (Humana Press, 2004).

184. Kruger, N. J. *The Bradford Method for Protein Quantitation* (Humana Press, 2002).

185. Yuan, X., Russell, T., Wood, G. & Desiderio, D. M. Analysis of the human lumbar cerebrospinal fluid proteome. *Electrophoresis* **23**, 1185–96 (2002). URL http://www.ncbi.nlm.nih.gov/pubmed/11981868.

186. Duan, X. *et al.* A straightforward and highly efficient precipitation/on-pellet digestion procedure coupled with a long gradient nano-LC separation and Orbitrap mass spectrometry for label-free expression profiling of the swine heart mitochondrial proteome. *Journal of proteome research* **8**, 2838–50 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/19290621http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2734143.

187. Sickmann, A. *et al.* Identification of proteins from human cerebrospinal fluid, separated by two-dimensional polyacrylamide gel electrophoresis. *Electrophoresis* **21**, 2721–2728 (2000). URL http://www.ncbi.nlm.nih.gov/pubmed/10949151http://doi.wiley.com/10.1002/1522-2683{%}2820000701{%}2921{%}3A13{%}3C2721{%}3A{%}3AAID-ELPS2721{%}3E3.0.CO{%}3B2-3.

188. Duarte, T. T. & Spencer, C. T. Personalized Proteomics: The Future of Precision Medicine. *Proteomes* **4** (2016). URL http://www.ncbi.nlm.nih.gov/pubmed/27882306http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5117667.

189. Gleave, M. E., Hsieh, J. T., von Eschenbach, A. C. & Chung, L. W. Prostate and bone fibroblasts induce human prostate cancer growth in vivo: implications for bidirectional tumor-stromal cell interaction in prostate carcinoma growth and metastasis. *The Journal of urology* **147**, 1151–9 (1992). URL http://www.ncbi.nlm.nih.gov/pubmed/1372662.

190. Gleave, M., Hsieh, J. T., Gao, C. A., von Eschenbach, A. C. & Chung, L. W. Acceleration of human prostate cancer growth in vivo by factors produced

by prostate and bone fibroblasts. *Cancer research* **51**, 3753–61 (1991). URL http://www.ncbi.nlm.nih.gov/pubmed/1712249.

191. Kaighn, M. E., Narayan, K. S., Ohnuki, Y., Lechner, J. F. & Jones, L. W. Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Investigative urology* **17**, 16–23 (1979). URL http://www.ncbi.nlm.nih.gov/pubmed/447482.

192. Berthon, P., Cussenot, O., Hopwood, L., Leduc, A. & Maitland, N. Functional expression of sv40 in normal human prostatic epithelial and fibroblastic cells - differentiation pattern of nontumorigenic cell-lines. *International journal of oncology* **6**, 333–43 (1995). URL http://www.ncbi.nlm.nih.gov/pubmed/21556542.

193. Dozmorov, M. G. *et al.* Unique patterns of molecular profiling between human prostate cancer LNCaP and PC-3 cells. *The Prostate* **69**, 1077–90 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/19343732http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2755240.

194. DePrimo, S. E. *et al.* Transcriptional programs activated by exposure of human prostate cancer cells to androgen. *Genome biology* **3**, RESEARCH0032 (2002). URL http://www.ncbi.nlm.nih.gov/pubmed/12184806http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC126237.

195. Solaini, G., Sgarbi, G. & Baracca, A. Oxidative phosphorylation in cancer cells. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1807**, 534–542 (2011).

196. Ramteke, A. *et al.* Exosomes secreted under hypoxia enhance invasiveness and stemness of prostate cancer cells by targeting adherens junction molecules. *Molecular Carcinogenesis* **54**, 554–565 (2015). URL http://doi.wiley.com/10.1002/mc.22124.

197. Knights, A. J., Funnell, A. P. W., Crossley, M. & Pearson, R. C. M. Holding Tight: Cell Junctions and Cancer Spread. *Trends in cancer research* **8**, 61–69 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/23450077http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3582402.

198. Hanahan, D. & Weinberg, R. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).

199. Su, L., Hershberger, R. J. & Weissman, I. L. LYAR, a novel nucleolar protein with zinc finger DNA-binding motifs, is involved in cell growth regulation.

*Genes & development* **7**, 735–48 (1993). URL http://www.ncbi.nlm.nih.gov/pubmed/8491376.

200. Wei, H., Mundade, R., Lange, K. C. & Lu, T. Protein arginine methylation of non-histone proteins and its role in diseases. *Cell cycle (Georgetown, Tex.)* **13**, 32–41 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/24296620http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3925732.

201. Prescott, J. L. & Tindall, D. J. Clathrin Gene Expression Is Androgen Regulated in the Prostate <sup>1</sup>. *Endocrinology* **139**, 2111–2119 (1998). URL http://www.ncbi.nlm.nih.gov/pubmed/9529000http://press.endocrine.org/doi/10.1210/endo.139.4.5926.

202. Kiessling, A. *et al.* Expression, regulation and function of the ISGylation system in prostate cancer. *Oncogene* **28**, 2606–2620 (2009). URL http://www.nature.com/doifinder/10.1038/onc.2009.115.

203. Chakraborty, P. K. *et al.* G ProteinCoupled Receptor Kinase GRK5 Phosphorylates Moesin and Regulates Metastasis in Prostate Cancer. *Cancer Research* **74** (2014).

204. Khan, A. P. *et al.* Quantitative proteomic profiling of prostate cancer reveals a role for miR-128 in prostate cancer. *Molecular & cellular proteomics : MCP* **9**, 298–312 (2010). URL http://www.ncbi.nlm.nih.gov/pubmed/19955085http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2830841.

205. Lin, L. *et al.* Transcriptional regulation of STAT3 by SPTBN1 and SMAD3 in HCC through cAMP-response element-binding proteins ATF3 and CREB2. *Carcinogenesis* **35**, 2393–403 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/25096061.

206. Wakamatsu, A. *et al.* Analysis of splicing variants of human full-length cDNAs. *DNA Research* (2009).

207. Leussis, M. P. *et al.* Ankyrin 3: genetic association with bipolar disorder and relevance to disease pathophysiology. *Biology of Mood & Anxiety Disorders* **2**, 18 (2012). URL http://biolmoodanxietydisord.biomedcentral.com/articles/10.1186/2045-5380-2-18.

# Addenda

## Supplementary tables

Table S1: Proteins identified during tandem mass spectrometry of the LNCaP cell line whole cell samples containing extracellular or membrane bound GO term annotations obtained from the NCBI database that may act as candidate biomarkers for PCa.

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| P31150 | GDIA_HUMAN | Rab GDP dissociation inhibitor alpha (Rab GDI alpha) (Guanosine diphosphate dissociation inhibitor 1) (GDI-1) (Oligophrenin-2) (Protein XAP-4) | *GDI1 GDIL OPHN2 RABGDIA XAP4* |
| P50851 | LRBA_HUMAN | Lipopolysaccharide-responsive and beige-like anchor protein (Beige-like protein) (CDC4-like protein) | *LRBA BGL CDC4L LBA* |
| P51688 | SPHM_HUMAN | N-sulphoglucosamine sulphohydrolase (EC 3.10.1.1) (Sulfoglucosamine sulfamidase) (Sulphamidase) | *SGSH HSS* |
| P53367 | ARFP1_HUMAN | Arfaptin-1 (ADP-ribosylation factor-interacting protein 1) | *ARFIP1* |
| P53985 | MOT1_HUMAN | Monocarboxylate transporter 1 (MCT 1) (Solute carrier family 16 member 1) | *SLC16A1 MCT1* |
| P54802 | ANAG_HUMAN | Alpha-N-acetylglucosaminidase (EC 3.2.1.50) (N-acetyl-alpha-glucosaminidase) (NAG) [Cleaved into: Alpha-N-acetylglucosaminidase 82 kDa form; Alpha-N-acetylglucosaminidase 77 kDa form] | *NAGLU UFHSD1* |
| P54886 | P5CS_HUMAN | Delta-1-pyrroline-5-carboxylate synthase (P5CS) (Aldehyde dehydrogenase family 18 member A1) [Includes: Glutamate 5-kinase (GK) (EC 2.7.2.11) (Gamma-glutamyl kinase); Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde dehydrogenase)] | *ALDH18A1 GSAS P5CS PYCS* |
| P56556 | NDUA6_HUMAN | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 6 (Complex I-B14) (CI-B14) (LYR motif-containing protein 6) (NADH-ubiquinone oxidoreductase B14 subunit) | *NDUFA6 LYRM6 NADHB14* |
| P78316 | NOP14_HUMAN | Nucleolar protein 14 (Nucleolar complex protein 14) | *NOP14 C4orf9 NOL14 RES4-25* |
| P82932 | RT06_HUMAN | 28S ribosomal protein S6, mitochondrial (MRP-S6) (S6mt) | *MRPS6 C21orf101 RPMS6* |

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| Q01831 | XPC_HUMAN | DNA repair protein complementing XP-C cells (Xeroderma pigmentosum group C-complementing protein) (p125) | *XPC XPCC* |
| Q02127 | PYRD_HUMAN | Dihydroorotate dehydrogenase (quinone), mitochondrial (DHOdehase) (EC 1.3.5.2) (Dihydroorotate oxidase) | *DHODH* |
| Q05682 | CALD1_HUMAN | Caldesmon (CDM) | *CALD1 CAD CDM* |
| Q12955 | ANK3_HUMAN | Ankyrin-3 (ANK-3) (Ankyrin-G) | *ANK3* |
| Q13505 | MTX1_HUMAN | Metaxin-1 (Mitochondrial outer membrane import complex protein 1) | *MTX1 MTX MTXN* |
| Q13823 | NOG2_HUMAN | Nucleolar GTP-binding protein 2 (Autoantigen NGP-1) | *GNL2 NGP1* |
| Q14353 | GAMT_HUMAN | Guanidinoacetate N-methyltransferase (EC 2.1.1.2) | *GAMT* |
| Q16134 | ETFD_HUMAN | Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial (ETF-QO) (ETF-ubiquinone oxidoreductase) (EC 1.5.5.1) (Electron-transferring-flavoprotein dehydrogenase) (ETF dehydrogenase) | *ETFDH* |
| Q16795 | NDUA9_HUMAN | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial (Complex I-39kD) (CI-39kD) (NADH-ubiquinone oxidoreductase 39 kDa subunit) | *NDUFA9 NDUFS2L* |
| Q16822 | PCKGM_HUMAN | Phosphoenolpyruvate carboxykinase [GTP], mitochondrial (PEPCK-M) (EC 4.1.1.32) | *PCK2 PEPCK2* |
| Q5TH69 | BIG3_HUMAN | Brefeldin A-inhibited guanine nucleotide-exchange protein 3 (ARFGEF family member 3) | *ARFGEF3 BIG3 C6orf92 KIAA1244* |
| Q5VZR0 | Q5VZR0_HUMAN | Golgi-associated plant pathogenesis-related protein 1 | *GLIPR2* |
| Q6YN16 | HSDL2_HUMAN | Hydroxysteroid dehydrogenase-like protein 2 (EC 1.-.-.-) (Short chain dehydrogenase/reductase family 13C member 1) | *HSDL2 C9orf99 SDR13C1* |
| Q8N490 | PNKD_HUMAN | Probable hydrolase PNKD (EC 3.-.-.-) (Myofibrillogenesis regulator 1) (MR-1) (Paroxysmal nonkinesiogenic dyskinesia protein) (Trans-activated by hepatitis C virus core protein 2) | *PNKD KIAA1184 MR1 TAHCCP2 FKSG19 UNQ2491 PRO5778* |
| Q8NE62 | CHDH_HUMAN | Choline dehydrogenase, mitochondrial (CDH) (CHD) (EC 1.1.99.1) | *CHDH* |
| Q8TCC3 | RM30_HUMAN | 39S ribosomal protein L30, mitochondrial (L30mt) (MRP-L30) (39S ribosomal protein L28, mitochondrial) (L28mt) (MRP-L28) | *MRPL30 MRPL28 RPML28 HSPC249* |
| Q92506 | DHB8_HUMAN | Estradiol 17-beta-dehydrogenase 8 (EC 1.1.1.62) (17-beta-hydroxysteroid dehydrogenase 8) (17-beta-HSD 8) (3-ketoacyl-[acyl-carrier-protein] reductase alpha subunit) (KAR alpha subunit) (3-oxoacyl-[acyl-carrier-protein] reductase) (EC 1.1.1.-) (Protein Ke6) (Ke-6) (Really interesting new gene 2 protein) (Short chain dehydrogenase/reductase family 30C member 1) (Testosterone 17-beta-dehydrogenase 8) (EC 1.1.1.239) | *HSD17B8 FABGL HKE6 RING2 SDR30C1* |
| Q969X5 | ERGI1_HUMAN | Endoplasmic reticulum-Golgi intermediate compartment protein 1 (ER-Golgi intermediate compartment 32 kDa protein) (ERGIC-32) | *ERGIC1 ERGIC32 KIAA1181 HT034* |
| Q96DV4 | RM38_HUMAN | 39S ribosomal protein L38, mitochondrial (L38mt) (MRP-L38) | *MRPL38 HSPC262* |
| Q96ED9 | HOOK2_HUMAN | Protein Hook homolog 2 (h-hook2) (hHK2) | *HOOK2* |

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| Q96L92 | SNX27_HUMAN | Sorting nexin-27 | *SNX27 KIAA0488 My014* |
| Q99543 | DNJC2_HUMAN | DnaJ homolog subfamily C member 2 (M-phase phosphoprotein 11) (Zuotin-related factor 1) [Cleaved into: DnaJ homolog subfamily C member 2, N-terminally processed] | *DNAJC2 MPHOSPH11 MPP11 ZRF1* |
| Q99988 | GDF15_HUMAN | Growth/differentiation factor 15 (GDF-15) (Macrophage inhibitory cytokine 1) (MIC-1) (NSAID-activated gene 1 protein) (NAG-1) (NSAID-regulated gene 1 protein) (NRG-1) (Placental TGF-beta) (Placental bone morphogenetic protein) (Prostate differentiation factor) | *GDF15 MIC1 PDF PLAB PTGFB* |
| Q9BYC8 | RM32_HUMAN | 39S ribosomal protein L32, mitochondrial (L32mt) (MRP-L32) | *MRPL32 HSPC283* |
| Q9BYC9 | RM20_HUMAN | 39S ribosomal protein L20, mitochondrial (L20mt) (MRP-L20) | *MRPL20* |
| Q9BYD1 | RM13_HUMAN | 39S ribosomal protein L13, mitochondrial (L13mt) (MRP-L13) | *MRPL13* |
| Q9BYN8 | RT26_HUMAN | 28S ribosomal protein S26, mitochondrial (MRP-S26) (S26mt) (28S ribosomal protein S13, mitochondrial) (MRP-S13) (S13mt) | *MRPS26 C20orf193 RPMS13* |
| Q9BZE1 | RM37_HUMAN | 39S ribosomal protein L37, mitochondrial (L37mt) (MRP-L37) (39S ribosomal protein L2, mitochondrial) (L2mt) (MRP-L2) | *MRPL37 MRPL2 RPML2 HSPC235* |
| Q9H8Y5 | ANKZ1_HUMAN | Ankyrin repeat and zinc finger domain-containing protein 1 (Zinc finger protein 744) | *ANKZF1 ZNF744* |
| Q9H9J2 | RM44_HUMAN | 39S ribosomal protein L44, mitochondrial (L44mt) (MRP-L44) (EC 3.1.26.-) | *MRPL44* |
| Q9NPJ3 | ACO13_HUMAN | Acyl-coenzyme A thioesterase 13 (Acyl-CoA thioesterase 13) (EC 3.1.2.-) (Thioesterase superfamily member 2) [Cleaved into: Acyl-coenzyme A thioesterase 13, N-terminally processed] | *ACOT13 THEM2 HT012 PNAS-27* |
| Q9NQW7 | XPP1_HUMAN | Xaa-Pro aminopeptidase 1 (EC 3.4.11.9) (Aminoacylproline aminopeptidase) (Cytosolic aminopeptidase P) (Soluble aminopeptidase P) (sAmp) (X-Pro aminopeptidase 1) (X-prolyl aminopeptidase 1, soluble) | *XPNPEP1 XPNPEPL XPN-PEPL1* |
| Q9NX20 | RM16_HUMAN | 39S ribosomal protein L16, mitochondrial (L16mt) (MRP-L16) | *MRPL16 PNAS-111* |
| Q9NY61 | AATF_HUMAN | Protein AATF (Apoptosis-antagonizing transcription factor) (Rb-binding protein Che-1) | *AATF CHE1 DED HSPC277* |
| Q9UHY7 | ENOPH_HUMAN | Enolase-phosphatase E1 (EC 3.1.3.77) (2,3-diketo-5-methylthio-1-phosphopentane phosphatase) (MASA homolog) | *ENOPH1 MASA MSTP145* |
| Q9UI09 | NDUAC_HUMAN | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12 (13 kDa differentiation-associated protein) (Complex I-B17.2) (CI-B17.2) (CIB17.2) (NADH-ubiquinone oxidoreductase subunit B17.2) | *NDUFA12 DAP13* |
| Q9UMX0 | UBQL1_HUMAN | Ubiquilin-1 (Protein linking IAP with cytoskeleton 1) (PLIC-1) (hPLIC-1) | *UBQLN1 DA41 PLIC1* |
| Q9Y244 | POMP_HUMAN | Proteasome maturation protein (Proteassemblin) (Protein UMP1 homolog) (hUMP1) (Voltage-gated K channel beta subunit 4.1) | *POMP C13orf12 UMP1 HSPC014 HSPC036 PNAS-110* |

Table S2: Proteins identified during tandem mass spectrometry of the PC-3 cell line whole cell samples containing extracellular or membrane bound GO term annotations obtained from the NCBI database that may act as candidate biomarkers for PCa. Protein and gene information was obtained from the UniProt database.

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| E7ESA6 | E7ESA6_HUMAN | Focal adhesion kinase 1 | *PTK2* |
| E7EV59 | E7EV59_HUMAN | V-type proton ATPase subunit C 1 | *ATP6V1C1* |
| E9PGC5 | E9PGC5_HUMAN | Receptor-type tyrosine-protein phosphatase kappa | *PTPRK* |
| F5GX39 | F5GX39_HUMAN | Transmembrane emp24 domain-containing protein 2 | *TMED2* |
| F5H442 | F5H442_HUMAN | Tumour susceptibility gene 101 protein | *TSG101* |
| F5H459 | F5H459_HUMAN | AP-3 complex subunit sigma-1 | *AP3S1* |
| G3XAL9 | G3XAL9_HUMAN | Solute carrier family 12 (Sodium/potassium/chloride transporters), member 2, isoform CRA_a (Solute carrier family 12 member 2) | *SLC12A2* *hCG_27034* |
| H0YCN4 | H0YCN4_HUMAN | DCN1-like protein (Defective in cullin neddylation protein 1-like protein) (Fragment) | *DCUN1D5* |
| H0YDT8 | H0YDT8_HUMAN | ER membrane protein complex subunit 7 (Fragment) | *EMC7* |
| H3BRB1 | H3BRB1_HUMAN | Uncharacterized protein | N/A |
| K7EL21 | K7EL21_HUMAN | Calcyphosin (Calcyphosine, isoform CRA_b) | *CAPS* *hCG_23051* |
| O00178 | GTPB1_HUMAN | GTP-binding protein 1 (G-protein 1) (GP-1) (GP1) | *GTPBP1* |
| O00231 | PSD11_HUMAN | 26S proteasome non-ATPase regulatory subunit 11 (26S proteasome regulatory subunit RPN6) (26S proteasome regulatory subunit S9) (26S proteasome regulatory subunit p44.5) | *PSMD11* |
| O14613 | BORG1_HUMAN | Cdc42 effector protein 2 (Binder of Rho GTPases 1) | *CDC42EP2* *BORG1* *CEP2* |
| O15145 | ARPC3_HUMAN | Actin-related protein 2/3 complex subunit 3 (Arp2/3 complex 21 kDa subunit) (p21-ARC) | *ARPC3* *ARC21* |
| O43491 | E41L2_HUMAN | Band 4.1-like protein 2 (Generally expressed protein 4.1) (4.1G) | *EPB41L2* |
| O75116 | ROCK2_HUMAN | Rho-associated protein kinase 2 (EC 2.7.11.1) (Rho kinase 2) (Rho-associated, coiled-coil-containing protein kinase 2) (Rho-associated, coiled-coil-containing protein kinase II) (ROCK-II) (p164 ROCK-2) | *ROCK2* *KIAA0619* |
| P01034 | CYTC_HUMAN | Cystatin-C (Cystatin-3) (Gamma-trace) (Neuroendocrine basic polypeptide) (Post-gamma-globulin) | *CST3* |
| P01699 | LV144_HUMAN | Immunoglobulin lambda variable 1-44 (Ig lambda chain V-I region MEM) (Ig lambda chain V-I region VOR) | *IGLV1-44* |
| P05556 | ITB1_HUMAN | Integrin beta-1 (Fibronectin receptor subunit beta) (Glycoprotein IIa) (GPIIA) (VLA-4 subunit beta) (CD antigen CD29) | *ITGB1* *FNRB* *MDF2* *MSK12* |
| P06576 | ATPB_HUMAN | ATP synthase subunit beta, mitochondrial (EC 3.6.3.14) | *ATP5B* *ATPMB* *ATPSB* |
| P08574 | CY1_HUMAN | Cytochrome c1, heme protein, mitochondrial (Complex III subunit 4) (Complex III subunit IV) (Cytochrome b-c1 complex subunit 4) (Ubiquinol-cytochrome-c reductase complex cytochrome c1 subunit) (Cytochrome c-1) | *CYC1* |
| P09467 | F16P1_HUMAN | Fructose-1,6-bisphosphatase 1 (FBPase 1) (EC 3.1.3.11) (D-fructose-1,6-bisphosphate 1-phosphohydrolase 1) (Liver FBPase) | *FBP1* *FBP* |
| P15090 | FABP4_HUMAN | Fatty acid-binding protein, adipocyte (Adipocyte lipid-binding protein) (ALBP) (Adipocyte-type fatty acid-binding protein) (A-FABP) (AFABP) (Fatty acid-binding protein 4) | *FABP4* |

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| P15407 | FOSL1_HUMAN | Fos-related antigen 1 (FRA-1) | *FOSL1 FRA1* |
| P15428 | PGDH_HUMAN | 15-hydroxyprostaglandin dehydrogenase [NAD(+)] (15-PGDH) (EC 1.1.1.141) (Prostaglandin dehydrogenase 1) (Short chain dehydrogenase/reductase family 36C member 1) | *HPGD PGDH1 SDR36C1* |
| P22681 | CBL_HUMAN | E3 ubiquitin-protein ligase CBL (EC 6.3.2.-) (Casitas B-lineage lymphoma proto-oncogene) (Proto-oncogene c-Cbl) (RING finger protein 55) (Signal transduction protein CBL) | *CBL CBL2 RNF55* |
| P24539 | AT5F1_HUMAN | ATP synthase F(0) complex subunit B1, mitochondrial (ATP synthase proton-transporting mitochondrial F(0) complex subunit B1) (ATP synthase subunit b) (ATPase subunit b) | *ATP5F1* |
| P40123 | CAP2_HUMAN | Adenylyl cyclase-associated protein 2 (CAP 2) | *CAP2* |
| P45877 | PPIC_HUMAN | Peptidyl-prolyl cis-trans isomerase C (PPIase C) (EC 5.2.1.8) (Cyclophilin C) (Rotamase C) | *PPIC CYPC* |
| P49841 | GSK3B_HUMAN | Glycogen synthase kinase-3 beta (GSK-3 beta) (EC 2.7.11.26) (Serine/threonine-protein kinase GSK3B) (EC 2.7.11.1) | *GSK3B* |
| P50552 | VASP_HUMAN | Vasodilator-stimulated phosphoprotein (VASP) | *VASP* |
| P61586 | RHOA_HUMAN | Transforming protein RhoA (Rho cDNA clone 12) (h12) | *RHOA ARH12 ARHA RHO12* |
| P63208 | SKP1_HUMAN | S-phase kinase-associated protein 1 (Cyclin-A/CDK2-associated protein p19) (p19A) (Organ of Corti protein 2) (OCP-2) (Organ of Corti protein II) (OCP-II) (RNA polymerase II elongation factor-like protein) (SIII) (Transcription elongation factor B polypeptide 1-like) (p19skp1) | *SKP1 EMC19 OCP2 SKP1A TCEB1L* |
| Q00587 | BORG5_HUMAN | Cdc42 effector protein 1 (Binder of Rho GTPases 5) (Serum protein MSE55) | *CDC42EP1 BORG5 CEP1 MSE55* |
| Q13418 | ILK_HUMAN | Integrin-linked protein kinase (EC 2.7.11.1) (59 kDa serine/threonine-protein kinase) (ILK-1) (ILK-2) (p59ILK) | *ILK ILK1 ILK2* |
| Q14165 | MLEC_HUMAN | Malectin | *MLEC KIAA0152* |
| Q15124 | PGM5_HUMAN | Phosphoglucomutase-like protein 5 (Aciculin) (Phosphoglucomutase-related protein) (PGM-RP) | *PGM5 PGMRP* |
| Q1L6U9 | MSMP_HUMAN | Prostate-associated microseminoprotein (PC3-secreted microprotein) | *MSMP PSMP* |
| Q5JXI8 | Q5JXI8_HUMAN | Four and a half LIM domains protein 1 (Fragment) | *FHL1* |
| Q5QPE1 | Q5QPE1_HUMAN | Protein transport protein Sec23B (Fragment) | *SEC23B* |
| Q8N8S7 | ENAH_HUMAN | Protein enabled homolog | *ENAH MENA* |
| Q8NI22 | MCFD2_HUMAN | Multiple coagulation factor deficiency protein 2 (Neural stem cell-derived neuronal survival protein) | *MCFD2 SDNSF* |
| Q8TAT6 | NPL4_HUMAN | Nuclear protein localization protein 4 homolog (Protein NPL4) | *NPLOC4 KIAA1499 NPL4* |
| Q92729 | PTPRU_HUMAN | Receptor-type tyrosine-protein phosphatase U (R-PTP-U) (EC 3.1.3.48) (Pancreatic carcinoma phosphatase 2) (PCP-2) (Protein-tyrosine phosphatase J) (PTP-J) (hPTP-J) (Protein-tyrosine phosphatase pi) (PTP pi) (Protein-tyrosine phosphatase receptor omicron) (PTP-RO) (Receptor-type protein-tyrosine phosphatase psi) (R-PTP-psi) | *PTPRU FMI PCP2 PTPRO* |
| Q96DG6 | CMBL_HUMAN | Carboxymethylenebutenolidase homolog (EC 3.1.-.-) | *CMBL* |
| Q96GD0 | PLPP_HUMAN | Pyridoxal phosphate phosphatase (PLP phosphatase) (EC 3.1.3.3) (EC 3.1.3.74) (Chronophin) | *PDXP CIN PLP PLPP* |
| Q99805 | TM9S2_HUMAN | Transmembrane 9 superfamily member 2 (p76) | *TM9SF2* |

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| Q9BS40 | LXN_HUMAN | Latexin (Endogenous carboxypeptidase inhibitor) (ECI) (Protein MUM) (Tissue carboxypeptidase inhibitor) (TCI) | *LXN* |
| Q9BXS5 | AP1M1_HUMAN | AP-1 complex subunit mu-1 (AP-mu chain family member mu1A) (Adaptor protein complex AP-1 subunit mu-1) (Adaptor-related protein complex 1 subunit mu-1) (Clathrin assembly protein complex 1 mu-1 medium chain 1) (Clathrin coat assembly protein AP47) (Clathrin coat-associated protein AP47) (Golgi adaptor HA1/AP1 adaptin mu-1 subunit) (Mu-adaptin 1) (Mu1A-adaptin) | *AP1M1* *CLTNM* |
| Q9H089 | LSG1_HUMAN | Large subunit GTPase 1 homolog (hLsg1) (EC 3.6.1.-) | *LSG1* |
| Q9H0U4 | RAB1B_HUMAN | Ras-related protein Rab-1B | *RAB1B* |
| Q9H299 | SH3L3_HUMAN | SH3 domain-binding glutamic acid-rich-like protein 3 (SH3 domain-binding protein 1) (SH3BP-1) | *SH3BGRL3* *P1725* |
| Q9NS15 | LTBP3_HUMAN | Latent-transforming growth factor beta-binding protein 3 (LTBP-3) | *LTBP3* |
| Q9NTJ5 | SAC1_HUMAN | Phosphatidylinositide phosphatase SAC1 (EC 3.1.3.-) (Suppressor of actin mutations 1-like protein) | *SACM1L* *KIAA0851* |
| Q9P289 | STK26_HUMAN | Serine/threonine-protein kinase 26 (EC 2.7.11.1) (MST3 and SOK1-related kinase) (Mammalian STE20-like protein kinase 4) (MST-4) (STE20-like kinase MST4) (Serine/threonine-protein kinase MASK) | *STK26* *MASK* *MST4* |
| Q9Y2T2 | AP3M1_HUMAN | AP-3 complex subunit mu-1 (AP-3 adaptor complex mu3A subunit) (Adaptor-related protein complex 3 subunit mu-1) (Mu-adaptin 3A) (Mu3A-adaptin) | *AP3M1* |
| Q9Y376 | CAB39_HUMAN | Calcium-binding protein 39 (MO25alpha) (Protein Mo25) | *CAB39* *MO25* *CGI-66* |
| Q9Y4G6 | TLN2_HUMAN | Talin-2 | *TLN2* *KIAA0320* |

Table S3: Proteins identified during tandem mass spectrometry of the LNCaP and PC-3 cell lines whole cell samples containing extracellular or membrane bound GO term annotations obtained from the NCBI database that may act as candidate biomarkers for PCa. Protein and gene information was obtained from the UniProt database.

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| C9JI87 | C9JI87_HUMAN | Voltage-dependent anion-selective channel protein 1 (Fragment) | *VDAC1* |
| E7EVA0 | E7EVA0_HUMAN | Microtubule-associated protein | *MAP4* |
| F8W7Q4 | F8W7Q4_HUMAN | Protein FAM162A | *FAM162A* |
| O43633 | CHM2A_HUMAN | Charged multivesicular body protein 2a (Chromatin-modifying protein 2a) (CHMP2a) (Putative breast adenocarcinoma marker BC-2) (Vacuolar protein sorting-associated protein 2-1) (Vps2-1) (hVps2-1) | *CHMP2A* *BC2* *CHMP2* |
| O75521 | ECI2_HUMAN | Enoyl-CoA delta isomerase 2, mitochondrial (EC 5.3.3.8) (DRS-1) (Delta(3),delta(2)-enoyl-CoA isomerase) (D3,D2-enoyl-CoA isomerase) (Diazepam-binding inhibitor-related protein 1) (DBI-related protein 1) (Dodecenoyl-CoA isomerase) (Hepatocellular carcinoma-associated antigen 88) (Peroxisomal 3,2-trans-enoyl-CoA isomerase) (pECI) (Renal carcinoma antigen NY-REN-1) | *ECI2* *DRS1* *HCA88* *PECI* |
| O95202 | LETM1_HUMAN | LETM1 and EF-hand domain-containing protein 1, mitochondrial (Leucine zipper-EF-hand-containing transmembrane protein 1) | *LETM1* |

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| O95785 | WIZ_HUMAN | Protein Wiz (Widely-interspaced zinc finger-containing protein) (Zinc finger protein 803) | *WIZ ZNF803* |
| P00167 | CYB5_HUMAN | Cytochrome b5 (Microsomal cytochrome b5 type A) (MCB5) | *CYB5A CYB5* |
| P01597 | KV139_HUMAN | Immunoglobulin kappa variable 1-39 (Ig kappa chain V-I region DEE) (Ig kappa chain V-I region Hau) (Ig kappa chain V-I region Mev) (Ig kappa chain V-I region OU) (Ig kappa chain V-I region Walker) | *IGKV1-39* |
| P04350 | TBB4A_HUMAN | Tubulin beta-4A chain (Tubulin 5 beta) (Tubulin beta-4 chain) | *TUBB4A TUBB4 TUBB5* |
| P08133 | ANXA6_HUMAN | Annexin A6 (67 kDa calelectrin) (Annexin VI) (Annexin-6) (Calphobindin-II) (CPB-II) (Chromobindin-20) (Lipocortin VI) (Protein III) (p68) (p70) | *ANXA6 ANX6* |
| P10253 | LYAG_HUMAN | Lysosomal alpha-glucosidase (EC 3.2.1.20) (Acid maltase) (Aglucosidase alfa) [Cleaved into: 76 kDa lysosomal alpha-glucosidase; 70 kDa lysosomal alpha-glucosidase] | *GAA* |
| P12004 | PCNA_HUMAN | Proliferating cell nuclear antigen (PCNA) (Cyclin) | *PCNA* |
| P20618 | PSB1_HUMAN | Proteasome subunit beta type-1 (EC 3.4.25.1) (Macropain subunit C5) (Multicatalytic endopeptidase complex subunit C5) (Proteasome component C5) (Proteasome gamma chain) | *PSMB1 PSC5* |
| P27635 | RL10_HUMAN | 60S ribosomal protein L10 (Laminin receptor homolog) (Protein QM) (Tumour suppressor QM) | *RPL10 DXS648E QM* |
| P28482 | MK01_HUMAN | Mitogen-activated protein kinase 1 (MAP kinase 1) (MAPK 1) (EC 2.7.11.24) (ERT1) (Extracellular signal-regulated kinase 2) (ERK-2) (MAP kinase isoform p42) (p42-MAPK) (Mitogen-activated protein kinase 2) (MAP kinase 2) (MAPK 2) | *MAPK1 ERK2 PRKM1 PRKM2* |
| P32119 | PRDX2_HUMAN | Peroxiredoxin-2 (EC 1.11.1.15) (Natural killer cell-enhancing factor B) (NKEF-B) (PRP) (Thiol-specific antioxidant protein) (TSA) (Thioredoxin peroxidase 1) (Thioredoxin-dependent peroxide reductase 1) | *PRDX2 NKEFB TDPX1* |
| P35080 | PROF2_HUMAN | Profilin-2 (Profilin II) | *PFN2* |
| P40429 | RL13A_HUMAN | 60S ribosomal protein L13a (23 kDa highly basic protein) | *RPL13A* |
| P56181 | NDUV3_HUMAN | NADH dehydrogenase [ubiquinone] flavoprotein 3, mitochondrial (Complex I-9kD) (CI-9kD) (NADH-ubiquinone oxidoreductase 9 kDa subunit) (Renal carcinoma antigen NY-REN-4) | *NDUFV3* |
| P62942 | FKB1A_HUMAN | Peptidyl-prolyl cis-trans isomerase FKBP1A (PPIase FKBP1A) (EC 5.2.1.8) (12 kDa FK506-binding protein) (12 kDa FKBP) (FKBP-12) (Calstabin-1) (FK506-binding protein 1A) (FKBP-1A) (Immunophilin FKBP12) (Rotamase) | *FKBP1A FKBP1 FKBP12* |
| P69905 | HBA_HUMAN | Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain) | *HBA1; HBA2* |
| Q13228 | SBP1_HUMAN | Selenium-binding protein 1 (56 kDa selenium-binding protein) (SBP56) (SP56) | *SELENBP1 SBP* |
| Q13596 | SNX1_HUMAN | Sorting nexin-1 | *SNX1* |
| Q14527 | HLTF_HUMAN | Helicase-like transcription factor (EC 3.6.4.-) (EC 6.3.2.-) (DNA-binding protein/plasminogen activator inhibitor 1 regulator) (HIP116) (RING finger protein 80) (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 3) (Sucrose nonfermenting protein 2-like 3) | *HLTF HIP116A RNF80 SMARCA3 SNF2L3 ZBU1* |
| Q15833 | STXB2_HUMAN | Syntaxin-binding protein 2 (Protein unc-18 homolog 2) (Unc18-2) (Protein unc-18 homolog B) (Unc-18B) | *STXBP2 UNC18B* |
| Q5QNZ2 | Q5QNZ2_HUMAN | ATP synthase F(0) complex subunit B1, mitochondrial | *ATP5F1* |

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| Q96HS1 | PGAM5_HUMAN | Serine/threonine-protein phosphatase PGAM5, mitochondrial (EC 3.1.3.16) (Bcl-XL-binding protein v68) (Phosphoglycerate mutase family member 5) | *PGAM5* |
| Q99798 | ACON_HUMAN | Aconitate hydratase, mitochondrial (Aconitase) (EC 4.2.1.3) (Citrate hydro-lyase) | *ACO2* |
| Q9BV57 | MTND_HUMAN | 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase (EC 1.13.11.54) (Acireductone dioxygenase (Fe(2+)-requiring)) (ARD) (Fe-ARD) (Membrane-type 1 matrix metalloproteinase cytoplasmic tail-binding protein 1) (MTCBP-1) (Submergence-induced protein-like factor) (Sip-L) | *ADI1 MTCBP1 HMFT1638* |
| Q9UBS4 | DJB11_HUMAN | DnaJ homolog subfamily B member 11 (APOBEC1-binding protein 2) (ABBP-2) (DnaJ protein homolog 9) (ER-associated DNAJ) (ER-associated Hsp40 co-chaperone) (Endoplasmic reticulum DNA J domain-containing protein 3) (ER-resident protein ERdj3) (ERdj3) (ERj3p) (HEDJ) (Human DnaJ protein 9) (hDj-9) (PWP1-interacting protein 4) | *DNAJB11 EDJ ERJ3 HDJ9 PSEC0121 UNQ537 PRO1080* |
| Q9Y285 | SYFA_HUMAN | Phenylalanine–tRNA ligase alpha subunit (EC 6.1.1.20) (CML33) (Phenylalanyl-tRNA synthetase alpha subunit) (PheRS) | *FARSA FARS FARSL FARSLA* |
| Q9Y4L1 | HYOU1_HUMAN | Hypoxia up-regulated protein 1 (150 kDa oxygen-regulated protein) (ORP-150) (170 kDa glucose-regulated protein) (GRP-170) | *HYOU1 GRP170 ORP150* |
| Q9Y5J7 | TIM9_HUMAN | Mitochondrial import inner membrane translocase subunit Tim9 | *TIMM9 TIM9 TIM9A TIMM9A* |
| Q9Y5L4 | TIM13_HUMAN | Mitochondrial import inner membrane translocase subunit Tim13 | *TIMM13 TIM13B TIMM13A TIMM13B* |

Table S4: Extracellular and membrane bound proteins unique to the LNCaP cell line identified during mass spectrometry of preliminary secretome samples (n=1).

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| E9PDC3 | E9PDC3_HUMAN | Armadillo repeat protein deleted in velo-cardio-facial syndrome | *ARVCF* |
| P12830 | CADH1_HUMAN | Cadherin-1 (CAM 120/80) (Epithelial cadherin) (E-cadherin) (Uvomorulin) (CD antigen CD324) [Cleaved into: E-Cad/CTF1; E-Cad/CTF2; E-Cad/CTF3] | *CDH1 CDHE UVO* |
| P52209 | 6PGD_HUMAN | 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44) | *PGD PGDH* |
| P62714 | PP2AB_HUMAN | Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform (PP2A-beta) (EC 3.1.3.16) | *PPP2CB* |
| Q07954 | LRP1_HUMAN | Prolow-density lipoprotein receptor-related protein 1 (LRP-1) (Alpha-2-macroglobulin receptor) (A2MR) (Apolipoprotein E receptor) (APOER) (CD antigen CD91) [Cleaved into: Low-density lipoprotein receptor-related protein 1 85 kDa subunit (LRP-85); Low-density lipoprotein receptor-related protein 1 515 kDa subunit (LRP-515); Low-density lipoprotein receptor-related protein 1 intracellular domain (LRPICD)] | *LRP1 A2MR APR* |

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| Q15139 | KPCD1_HUMAN | Serine/threonine-protein kinase D1 (EC 2.7.11.13) (Protein kinase C mu type) (Protein kinase D) (nPKC-D1) (nPKC-mu) | *PRKD1* *PKD* *PKD1* *PRKCM* |
| Q3LXA3 | TKFC_HUMAN | Triokinase/FMN cyclase (Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase (cyclizing)) [Includes: ATP-dependent dihydroxyacetone kinase (DHA kinase) (EC 2.7.1.28) (EC 2.7.1.29) (Glycerone kinase) (Triokinase) (Triose kinase); FAD-AMP lyase (cyclizing) (EC 4.6.1.15) (FAD-AMP lyase (cyclic FMN forming)) (FMN cyclase)] | *TKFC* *DAK* |
| Q8WVM8 | SCFD1_HUMAN | Sec1 family domain-containing protein 1 (SLY1 homolog) (Sly1p) (Syntaxin-binding protein 1-like 2) | *SCFD1* *C14orf163* *KIAA0917* *STXBP1L2* *FKSG23* |

Table S5: Extracellular and membrane bound proteins unique to the PC-3 cell line identified during mass spectrometry of preliminary secretome samples (n=1).

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| E9PK47 | E9PK47_HUMAN | Alpha-1,4 glucan phosphorylase (EC 2.4.1.1) | *PYGL* |
| Q14533 | KRT81_HUMAN | Keratin, type II cuticular Hb1 (Hair keratin K2.9) (Keratin, hair, basic, 1) (Keratin-81) (K81) (Metastatic lymph node 137 gene protein) (MLN 137) (Type II hair keratin Hb1) (Type-II keratin Kb21) (ghHKb1) (ghHb1) | *KRT81* *KRTHB1* *MLN137* |
| Q96IU4 | ABHEB_HUMAN | Protein ABHD14B (EC 3.-.-.-) (Alpha/beta hydrolase domain-containing protein 14B) (Abhydrolase domain-containing protein 14B) (CCG1-interacting factor B) | *ABHD14B* *CIB* |

Table S6: Extracellular and membrane bound proteins common to the LNCaP and PC-3 cell lines identified during mass spectrometry of preliminary secretome samples (n=1).

| Entry | Entry name | Protein names | Gene names |
|---|---|---|---|
| O43396 | TXNL1_HUMAN | Thioredoxin-like protein 1 (32 kDa thioredoxin-related protein) | *TXNL1* *TRP32* *TXL* *TXNL* |
| P21291 | CSRP1_HUMAN | Cysteine and glycine-rich protein 1 (Cysteine-rich protein 1) (CRP) (CRP1) (Epididymis luminal protein 141) (HEL-141) | *CSRP1* *CSRP* *CYRP* |
| P84077 | ARF1_HUMAN | ADP-ribosylation factor 1 | *ARF1* |
| Q13200 | PSMD2_HUMAN | 26S proteasome non-ATPase regulatory subunit 2 (26S proteasome regulatory subunit RPN1) (26S proteasome regulatory subunit S2) (26S proteasome subunit p97) (Protein 55.11) (Tumour necrosis factor type 1 receptor-associated protein 2) | *PSMD2* *TRAP2* |